

From: Chan, Christina
Sent: Wednesday, January 26, 2005 2:45 PM
To: Ramirez, Delia; STIC-Biotech/ChemLib
Subject: RE: rush search 09/886400

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

CRFE

-----Original Message-----

From: Ramirez, Delia
Sent: Wednesday, January 26, 2005 2:39 PM
To: Chan, Christina
Subject: rush search 09/886400

Hi Christina,

Please approve the following interference search: SEQ ID NO: 4 in the protein and nucleic acid databases

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

STAFF USE ONLY

Searcher: [Signature]
Searcher Phone: 2-2504
Date Searcher Picked up: 1/27/05
Date Completed: 2/1/05
Searcher Prep/Rev. Time: 10
Online Time: 15

Type of Search

NA Sequence: # ✓
AA Sequence: # ✓
Structure: # ✓
Bibliographic:
Litigation:
Patent Family:
Other:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2005, 05:58:26 ; Search time 62 seconds

(without alignments)
2121.118 Million cell updates/sec

Title: US-09-886-400A-4

Perfect score: 1879
Sequence: 1 LRAIVFHGNQYAIPEKSEI.....RLDAFRAIYNDWKGNGEPP 364

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289366 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|----------------------|
| 1 | 1879 | 100.0 | 364 | 14 | US-10-112-231A-4 |
| 2 | 1870 | 99.5 | 364 | 9 | US-09-886-400-4 |
| 3 | 1870 | 99.5 | 364 | 13 | US-10-112-357-4 |
| 4 | 1870 | 99.5 | 364 | 13 | US-10-114-403-4 |
| 5 | 1870 | 99.5 | 364 | 13 | US-10-116-606-4 |
| 6 | 1870 | 99.5 | 364 | 13 | US-10-112-331-4 |
| 7 | 1870 | 99.5 | 364 | 13 | US-10-112-377-4 |
| 8 | 1870 | 99.5 | 364 | 13 | US-10-116-581-4 |
| 9 | 1870 | 99.5 | 364 | 13 | US-10-112-442-4 |
| 10 | 1870 | 99.5 | 364 | 13 | US-10-112-418-4 |
| 11 | 1870 | 99.5 | 364 | 13 | US-10-114-083-4 |
| 12 | 111.5 | 5.9 | 890 | 15 | US-10-282-122A-53281 |
| 13 | 109 | 5.8 | 312 | 14 | US-10-369-493-23237 |

| | | | | | | |
|----|-------|-----|------|----|----------------------|-------------------|
| 14 | 106 | 5.6 | 590 | 17 | US-10-425-115-291380 | Sequence 291380, |
| 15 | 105.5 | 5.6 | 573 | 15 | US-10-424-599-175517 | Sequence 175517, |
| 16 | 105.5 | 5.6 | 730 | 15 | US-10-425-114-55544 | Sequence 55544, A |
| 17 | 103 | 5.5 | 1073 | 9 | US-09-815-242-12361 | Sequence 12361, A |
| 18 | 103 | 5.5 | 1147 | 9 | US-09-815-242-5468 | Sequence 5468, Ap |
| 19 | 103 | 5.5 | 1150 | 15 | US-10-282-122A-44391 | Sequence 44391, A |
| 20 | 102.5 | 5.5 | 478 | 14 | US-10-369-493-4506 | Sequence 4506, Ap |
| 21 | 102.5 | 5.5 | 478 | 14 | US-10-369-493-7264 | Sequence 7264, Ap |
| 22 | 101 | 5.4 | 441 | 16 | US-10-437-963-122721 | Sequence 122721, |
| 23 | 100.5 | 5.3 | 314 | 14 | US-10-369-493-17395 | Sequence 17395, A |
| 24 | 98.5 | 5.2 | 813 | 15 | US-10-282-122A-47208 | Sequence 47208, A |
| 25 | 98 | 5.2 | 510 | 14 | US-10-369-493-56 | Sequence 56, Appl |
| 26 | 97 | 5.2 | 652 | 15 | US-10-282-122A-63055 | Sequence 63055, A |
| 27 | 97 | 5.2 | 1878 | 15 | US-10-607-631-20 | Sequence 20, Appl |
| 28 | 96.5 | 5.1 | 798 | 15 | US-10-425-114-37809 | Sequence 37809, A |
| 29 | 96.5 | 5.1 | 808 | 17 | US-10-739-930-6793 | Sequence 6793, Ap |
| 30 | 96.5 | 5.1 | 1085 | 15 | US-10-282-122A-58833 | Sequence 58833, A |
| 31 | 96.5 | 5.1 | 3432 | 15 | US-10-282-122A-65849 | Sequence 65849, A |
| 32 | 96 | 5.1 | 530 | 14 | US-10-369-493-10268 | Sequence 10268, A |
| 33 | 96 | 5.1 | 530 | 15 | US-10-282-122A-77165 | Sequence 77165, A |
| 34 | 96 | 5.1 | 698 | 16 | US-10-437-963-122301 | Sequence 122301, |
| 35 | 95.5 | 5.1 | 875 | 15 | US-10-282-122A-52807 | Sequence 52807, A |
| 36 | 95 | 5.1 | 535 | 15 | US-10-282-122A-55460 | Sequence 55460, A |
| 37 | 95 | 5.1 | 2094 | 16 | US-10-437-963-161607 | Sequence 161607, |
| 38 | 94.5 | 5.0 | 749 | 14 | US-10-369-493-9015 | Sequence 9015, Ap |
| 39 | 94.5 | 5.0 | 1822 | 15 | US-10-432-443-39 | Sequence 39, Appl |
| 40 | 94 | 5.0 | 274 | 15 | US-10-282-122A-54582 | Sequence 54582, A |
| 41 | 94 | 5.0 | 664 | 14 | US-10-032-585-7195 | Sequence 7195, Ap |
| 42 | 93.5 | 5.0 | 993 | 14 | US-10-369-493-20014 | Sequence 20014, A |
| 43 | 93.5 | 5.0 | 1165 | 15 | US-10-282-122A-71768 | Sequence 71768, A |
| 44 | 93.5 | 5.0 | 1320 | 14 | US-10-161-051-167 | Sequence 167, App |
| 45 | 93.5 | 5.0 | 2049 | 16 | US-10-437-963-197248 | Sequence 197248, |

ALIGNMENTS

RESULT 1
US-10-112-231A-4
; Sequence 4, Application US/10112231A
; Publication No. US20030184276A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/112,231A
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 09/407,806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-231A-4

Query Match 100.0%; Score 1879; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 8,4e-166;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 LRAIVFHGNQYAIPEKSEIPIKVEKAYIPVETLKEEIPFGNITGYTLKPKQIID 60
Db 1 LRAIVFHGNQYAIPEKSEIPIKVEKAYIPVETLKEEIPFGNITGYTLKPKQIID 60
Cy 61 LVKGIASDLIEITGTSYTHAIPPLPLSRVAVQVQDRREVKEELFEVSPKGFVLPRLAY 120
Db 61 LVKGIASDLIEITGTSYTHAIPPLPLSRVAVQVQDRREVKEELFEVSPKGFVLPRLAY 120

```

QY 121 DPTIIPALIKDNGYEYLFPADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
DB 121 DPTIIPALIKDNGYEYLFADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
QY 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

RESULT 2

```

US-09-886-400-4
; Sequence 4, Application US/09886400
; Patent No. US20020045226A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-09-886-400-4

```

Query Match

```

Best Local Similarity 99.5%; Score 1870; DB 9; Length 364;
Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRALVFHGNLYOYAEIPKSEIPKVIKAYIPVETLKEBIPFGNITGYTLKFLPKDIID 60
DB 1 LRALVFHGNLYOYAEIPKSEIPKVIKAYIPVETLKEBIPFGNITGYTLKFLPKDIID 60
QY 61 LVKGGIADLIEITIGSYTHALPLPLSRVEAOVQOREVKEBELFEVSPKGFWLPBELAY 120
DB 61 LVKGGIADLIEITIGSYTHALPLPLSRVEAOVQOREVKEBELFEVSPKGFWLPBELAY 120
QY 121 DPTIIPALIKDNGYEYLFADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
DB 121 DPTIIPALIKDNGYEYLFADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
QY 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360

```

```

DB 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

RESULT 3

```

US-10-112-357-4
; Sequence 4, Application US/10112357
; Publication No. US20020115099A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,357
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-357-4

```

```

Query Match
Best Local Similarity 99.5%; Score 1870; DB 13; Length 364;
Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRALVFHGNLYOYAEIPKSEIPKVIKAYIPVETLKEBIPFGNITGYTLKFLPKDIID 60
DB 1 LRALVFHGNLYOYAEIPKSEIPKVIKAYIPVETLKEBIPFGNITGYTLKFLPKDIID 60
QY 61 LVKGGIADLIEITIGSYTHALPLPLSRVEAOVQOREVKEBELFEVSPKGFWLPBELAY 120
DB 61 LVKGGIADLIEITIGSYTHALPLPLSRVEAOVQOREVKEBELFEVSPKGFWLPBELAY 120
QY 121 DPTIIPALIKDNGYEYLFADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
DB 121 DPTIIPALIKDNGYEYLFADGEAMLPSAHLNSAIKPIKPYPHLIKAOREKFRYISYLLG 180
QY 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELKRAIKLVFEGKVTLLKAVKDIEAVPVWVAVNTAVMLGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRELYLTSSWAP 300
QY 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIWEDEGNARLNMLSYNMGELAFLAENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

```

RESULT 4

```

US-10-114-403-4
; Sequence 4, Application US/10114403
; Publication No. US20020115100A1
; GENERAL INFORMATION:

```

```
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/114,403
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PR
; ORGANISM: Thermococcus alcaliphilus
US-10-114-403-4
```

```
Query Match          99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LRALVFHGNLQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINTGYTLKFLPKDIIID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINTGYTLKFLPKDIIID 60
QY 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREKBELEFISPKGFWLPETLAY 120
DB 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREKBELEFISPKGFWLPETLAY 120
QY 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
DB 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
QY 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
DB 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
QY 181 LRELRAIKLVFEGSKVTLLKAVKDI EAVPVVAVTAVMLGIGRLPLNNPKVVASWIEDKD 240
DB 181 LRELRAIKLVFEGSKVTLLKAVKDI EAVPVVAVTAVMLGIGRLPLNNPKVVASWIEDKD 240
QY 241 NILLYGTDIEIFIGYRDIAGRMSVEGLLEVIDELNSCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEIFIGYRDIAGRMSVEGLLEVIDELNSCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIRREDEGNARLNLMLSYNMKGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
DB 301 DKSIRIRREDEGNARLNLMLSYNMKGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364
RESULT 5
US-10-116-606-4
; Sequence 4, Application US/10116606
; Publication No. US20020119515A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/116,606
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
```

```
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PR
; ORGANISM: Thermococcus alcaliphilus
US-10-116-606-4
```

```
Query Match          99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LRALVFHGNLQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINTGYTLKFLPKDIIID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVEKAYIPVETLKEEIPFGINTGYTLKFLPKDIIID 60
QY 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREKBELEFISPKGFWLPETLAY 120
DB 61 LVKGIASDLIEITIGSTYTHAILPLPLSRVEAOVQDRREKBELEFISPKGFWLPETLAY 120
QY 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
DB 121 DPTIPALIKONGYEYLPADGEAMLFSAHLSAIPKIPKLYPHLIKAKREKFRYISYLLG 180
QY 181 LRELRAIKLVFEGSKVTLLKAVKDI EAVPVVAVTAVMLGIGRLPLNNPKVVASWIEDKD 240
DB 181 LRELRAIKLVFEGSKVTLLKAVKDI EAVPVVAVTAVMLGIGRLPLNNPKVVASWIEDKD 240
QY 241 NILLYGTDIEIFIGYRDIAGRMSVEGLLEVIDELNSCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEIFIGYRDIAGRMSVEGLLEVIDELNSCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIRREDEGNARLNLMLSYNMKGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
DB 301 DKSIRIRREDEGNARLNLMLSYNMKGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
QY 361 NGEP 364
DB 361 NGEP 364
```

```
RESULT 6
US-10-112-331-4
; Sequence 4, Application US/10112331
; Publication No. US20020119550A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE 1
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,331
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PR
; ORGANISM: Thermococcus alcaliphilus
US-10-112-331-4
```

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
DB 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
QY 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIMEDEGNALNMLSYNMRGELAFLENSDARGMEPLPERRLDAPFAIYNDWGE 360
DB 301 DKSIRIMEDEGNALNMLSYNMRGELAFLENSDARGMEPLPERRLDAPFAIYNDWGE 360
QY 361 NGEF 364
DB 361 NGEF 364
```

RESULT 7

```
US-10-112-377-4
; Sequence 4, Application US/10112377
; Publication No. US20020120108A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-377-4
```

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
DB 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
```

```
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
QY 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIMEDEGNALNMLSYNMRGELAFLENSDARGMEPLPERRLDAPFAIYNDWGE 360
DB 301 DKSIRIMEDEGNALNMLSYNMRGELAFLENSDARGMEPLPERRLDAPFAIYNDWGE 360
QY 361 NGEF 364
DB 361 NGEF 364
```

RESULT 8

```
US-10-116-581-4
; Sequence 4, Application US/10116581
; Publication No. US20020137116A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVER1120-4
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-116-581-4
```

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5,7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
DB 1 LRALVFHGNQYVAEIPKSEIPKYTEKAYIPVETTLKEEIPFGINTGTTLKFLPXDID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKELFEVSPKGFMLPELAY 120
QY 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
DB 121 DPPIPALIKONGYEYLFADGEAMLFSAHLSAIPKIPLYPHILKAQREKFRFYSYLLG 180
QY 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
DB 181 LREIRKAIKLVFEGKVTLLKAVKDIEAVPVVAVTAVMLGIGRLPLMNPKKVASWIEDKD 240
QY 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILVGTDIIEFIIGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
```

```

QY 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

RESULT 9
US-10-112-442-4
; Sequence 4, Application US/10112442
; Publication No. US2002015097A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,442
; PRIOR APPLICATION NUMBER: 2002-03-29
; PRIOR FILING DATE: 2001-06-20
; PRIOR FILING DATE: 2001-06-20
; PRIOR FILING DATE: 2001-06-20
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-442-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPLPKDID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPLPKDID 60
QY 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFMLPELAY 120
QY 121 DPPIIPALIKONGYEYLPADGAMLFSAHNSAIKPIPLYPHLIKAKREKFRYISYLLG 180
DB 121 DPPIIPALIKONGYEYLPADGAMLFSAHNSAIKPIPLYPHLIKAKREKFRYISYLLG 180
QY 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLMPPKVAASWIEDKD 240
DB 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLMPPKVAASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELINSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELINSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

RESULT 10
US-10-112-418-4

```

```

; Sequence 4, Application US/10112418
; Publication No. US20020155486A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,418
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-418-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 5.7e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPLPKDID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGNITGYTLKPLPKDID 60
QY 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSTYTHAILPLPLSRVEAOVORDREYKEELFEVSPKGFMLPELAY 120
QY 121 DPPIIPALIKONGYEYLPADGAMLFSAHNSAIKPIPLYPHLIKAKREKFRYISYLLG 180
DB 121 DPPIIPALIKONGYEYLPADGAMLFSAHNSAIKPIPLYPHLIKAKREKFRYISYLLG 180
QY 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLMPPKVAASWIEDKD 240
DB 181 LRELKRAIKLVFEKSKVTLKAVKDIEAVPVVAVTAVMLGIGRLPLMPPKVAASWIEDKD 240
QY 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELINSELCLPSELKHSGRELYLRTSSWAP 300
DB 241 NILLYGTDIEFIIGYRDLAGYRMSVEGLLEVIDELINSELCLPSELKHSGRELYLRTSSWAP 300
QY 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
DB 301 DKSIRIRREDEGNARLMLSYNMRGELAFLENSDARGMEPLPERRLDAPRAIYNDWRGE 360
QY 361 NGEP 364
DB 361 NGEP 364

RESULT 11
US-10-114-083-4
; Sequence 4, Application US/10114083
; Publication No. US20020160464A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/114,083
; CURRENT FILING DATE: 2002-04-01

```

| | | | | |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match | 99.5% | Score 1870; | DB 13; | Length 364; |
| Best Local Similarity | 99.5%; | Pred. No. 5.7e-165; | | |
| Matches 362; | Conservative 1; | Mismatches 1; | Indels 0; | Gaps 0; |

RESULT 12
US-10-282

```

PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/270,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/230,335
PRIORITY FILING DATE: 2000-09-06
PRIORITY APPLICATION NUMBER: 60/230,347
PRIORITY FILING DATE: 2000-09-09
PRIORITY APPLICATION NUMBER: 60/242,558
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/267,636
PRIORITY FILING DATE: 2001-02-09
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 53281
LENGTH: 890
TYPE: PRT
ORGANISM: Clostridium difficile
US-10-282-122A-53281

```

| Query Match | 5.9% | Score 111.5 | DB 15 | Length 890 |
|-----------------------|-------|---------------------------------------------------------------|-------|------------|
| Best Local Similarity | 23.6% | Pred. No. 0.57 | 98 | Indels 87 |
| Matches | 72 | Conservative | 48 | Mismatches |
| QY | 53 | FLPKDIDLVKGGJASDLIEITGTSYTHAIPPLPLSFVEAOVORREKKELEFEVSPKG | 112 | |
| Db | 562 | FQP-DLIIAIVGGSGMDAGKIMWWWYEHPEVDFOLA-----NRFMIDIRKRV-VFPK | 612 | |
| QY | 113 | FWLPELAYDPIIP-----AII--KDNQYEYLFDQGEAMLPESAHNSAIKPIKP | 158 | |
| Db | 613 | --MGKAYFAAIPITSGAGSEVTPRAVITDODSGVKYPLADYELMPNMAIIDADMMEKP | 670 | |
| QY | 159 | LYPHL-----IKAQEKRFRTYSYL-----LGRLEPKAIKLFEE-----GKVTL | 198 | |
| Db | 671 | --PRLTAAGVDALTTHALEAVYSMLRTERPADGLA-----QACKIIFEYLPRAKYNGKNDK | 724 | |
| QY | 199 | KAVKDIEAVPVWAVNTA-----VMLGIGRLPLNN-----P | 229 | |
| Db | 725 | EAREKMAASTMAGSFPANFLGICSLAHKLGAHFHVOHGVANMLLINEVIFKNCAEAP | 784 | |
| QY | 230 | KKVASMIEDK--DNILLYGTDIETFGYDIAGY--RMSVEGLLEVIYIDENSLCLPSEAK | 285 | |
| Db | 785 | NKGAFSGYQRYPDCIORYA--EFASFAGIKGSTOEKVDNLIKAIIDELKAVGLPKPTIK | 841 | |
| QY | 286 | HSGRE | 290 | |
| Db | 842 | EAGVE | 846 | |

```

RESULT 13
US-10-369-493-23237
; Sequence 23237, Application US/10369493
; Publication NO. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

```



```

; SEQ ID NO 23237
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23237

```

| | | | | |
|-----------------------|------------------|----------------|------------|-------------|
| Query Match | 5.8% | Score 109; | DB 14; | Length 312; |
| Best Local Similarity | 23.2% | Pred No. 0.22; | | |
| Matches 44; | Conservative 32; | Mismatches 58; | Indels 56; | Gaps 9 |

```

QY 10 LQVAEIPKSIIP-----VLEKVIIVIELLKEELPEGLNIGTGLAEFKDIIID 61
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 130 MYIAVKESEFFPERIVIGOSVLDTR--PFTVAEE--LNLIS-----VKDVTCF 175
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 62 VKGGIASDLIEIIGTSYTHAI--LPPLPLSRVAQOVORDEVKEHPEVSEKQ--FWLPE 117
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 176 VLGGHGDWMLPVRYSYAGGIPELTIIPERIDALVERTRKGGGEIVNLLONGSAYAPA 235
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 118 LAYDPIIPALIKON-----GYEYLF-----ADGEAMLFIS-----A 147
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 236 ASLTEWEALIKQORVLPTIAVLEGEYGEISIGYVPTIVGNGLEQIIELETTYERA 285
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 148 HUNSAIKPIK 157
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 296 QLNKSVESVK 305
   ||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

RESULT 14
US-10-425-115-291380

```

/ Sequence 291380, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 291380
/ LENGTH: 590
/ TYPE: PRF
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1) ..(590)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_28825C.1.pep
/ US-10-425-115-291380

```

```
QY      224 LPIMNPKKVA-SW 235
      |  | : : | |
Db      568 LTQMDGTEVGNSW 580
```

```

RESULT 15
US-10-424-599-175517
Sequence: 175517, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ. ID NOS: 285684
SEQ ID NO 175517
LENGTH: 573
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_12950C.1.pep
US-10-424-599-175517

```

| | | | | | | | |
|-----------------------|--------------|-------|------------|----|--------|--------|------|
| Query Match | 5.64 | Score | 105.5 | DB | 15 | Length | 573 |
| Best Local Similarity | 21.34 | Pred. | No. 1.1 | | | | |
| Matches 77 | Conservative | 57 | Mismatches | 97 | Indels | 131 | Gaps |
| | | | | | | | 19 |

```

QY      6  PFGNLT-QVAEIPKRS-----EI PKYIEKAVIPV-----IETL 35
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      290  FLGSLGAKLGGVSSSLGSLGSEVPPVSPITPSFYCYWCMPGSLGCTPSLAIAVTQSPNSSTETL 349
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      36  IKEEIPFGINITYTLKFLPKDIIIDLVGKGIASDLIEIGTSTYHALPLPLSRVEAOV 95
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      350  ---PFGSGASLANPLS-----VNLT-----DPVQPLGISMDF--PLLP---DPLV 368
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      96  QRDREYKEELFEVSPKGFMLPELAYDP--IPAI-LKONGEYEIFADGEAMLFSAHLNSA 152
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      389  RMSLPTRSQCIPTFTF-----LMCDPIVHVPIVDVCSGCGYLVSAGMSPS----- 435
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      153  IKPIKLYPHLLKQAREKRFRIYISVLGLRRLKATILVEG-----KYLKAYKD 203
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      436  ---IPLEAHNLKPLIPSSDAV-----KARETLRLLSGSSGQNGQMMRPTLPPI-- 464
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      204  IEAVPVVAVVAANVAWMLGIGRLPLNMPKKVAMIEDKNIIILLYGTHDIEPIGYDI 257
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      485  -----LNTPD-----ENQNNILVAGSRGLYTRTRINNAIAMS 516
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      258  --AGYRMSVEGLLEVIDELNSELCLPSELKHSGLRELYLRTSSWAPDKSLRIWREDEGNAR 315
      1  : : : : : : : : : : : : : : : : : : : : : : : :
Db      517  IAAMGIVLSISGVSKVDGSVYSELC-----ENYGNILBAVKNSN--DSGGGAFLPDDEGGSS 568
      1  : : : : : : : : : : : : : : : : : : : : : : : :
QY      316  LN 317
      1  : :
Db      569  LD 570

```

Search completed: January 27, 2005, 05:59:41
Job time : 65 secs

This Page Blank (uspto)

Db 301 DKSRLRWEDEGNARLNLMLSYNMRGELAFLAENSDARGWEPLPERRLDAFRAIYNDWRGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 2
US-10-112-231A-4
; Sequence 4, Application US/10112231A
; Patent No. 6744246
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/112,231A
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 09/407,806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 08/613,220
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-231A-4

Query Match 100.0%; Score 1879; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4,9e-191;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LRAVFGHNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINITYTLKFLPKDIIID 60
Db 1 LRAVFGHNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINITYTLKFLPKDIIID 60
Qy 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVDAQVORDREYKEELFEVSPKGFMLPELAY 120
Db 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVDAQVORDREYKEELFEVSPKGFMLPELAY 120
Qy 121 DPPIPAIKDNGEYELFPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180
Db 121 DPPIPAIKDNGEYELFPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180
Qy 181 LRELKRAIKLVEEGKVTLLKAVKDIKAVVAVVAVTAVMLGIGRLPLMNPCKKVASWIEDKD 240
Db 181 LRELKRAIKLVEEGKVTLLKAVKDIKAVVAVVAVTAVMLGIGRLPLMNPCKKVASWIEDKD 240
Qy 241 NILLYGTDIEFIIGRDYAGYRMSVGLLEVIYDELINSELCLPSLKHSGRELYLRTSSWAP 300
Db 241 NILLYGTDIEFIIGRDYAGYRMSVGLLEVIYDELINSELCLPSLKHSGRELYLRTSSWAP 300
Qy 301 DKSRLRWEDEGNARLNLMLSYNMRGELAFLAENSDARGWEPLPERRLDAFRAIYNDWRGE 360
Db 301 DKSRLRWEDEGNARLNLMLSYNMRGELAFLAENSDARGWEPLPERRLDAFRAIYNDWRGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 3
US-08-613-220B-4
; Sequence 4, Application US/08613220B
; Patent No. 5958751
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson, P.C.
;; STREET: 4225 Executive Square, Suite 1400
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: US
;; ZIP: 92037

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows95
;; SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/613,220B
;; FILING DATE: 08-MAR-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hallie, Ph.D., Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 09010/004001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619-678-5070
;; TELEFAX: 619-68-5099

;; TELEX:
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 346 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear

;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: internal
US-08-613-220B-4

Query Match 84.6%; Score 1589; DB 2; Length 346;
Best Local Similarity 95.1%; Pred. No. 3.1e-160;
Matches 346; Conservative 0; Mismatches 0; Indels 18; Gaps 18;

Qy 1 LRAVFGHNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINITYTLKFLPKDIIID 60
Db 1 LRAVFGHNLQYAEIPKSEIPKVIKAYIPVETLKEEIPFGINITYTLKFLPKDIIID 57
Qy 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVDAQVORDREYKEELFEVSPKGFMLPELAY 120
Db 61 LVKGGIASDLIEITIGSTYTHAILPLPLSRVDAQVORDREYKEELFEVSPKGFMLPELAY 114
Qy 121 DPPIPAIKDNGEYELFPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 180
Db 121 DPPIPAIKDNGEYELFPADGEAMLFSALNSAIKPIKPLYPHLIKAKQREKFRYISYLLG 171
Qy 181 LRELKRAIKLVEEGKVTLLKAVKDIKAVVAVVAVTAVMLGIGRLPLMNPCKKVASWIEDKD 240
Db 181 LRELKRAIKLVEEGKVTLLKAVKDIKAVVAVVAVTAVMLGIGRLPLMNPCKKVASWIEDKD 228
Qy 241 NILLYGTDIEFIIGRDYAGYRMSVGLLEVIYDELINSELCLPSLKHSGRELYLRTSSWAP 300
Db 241 NILLYGTDIEFIIGRDYAGYRMSVGLLEVIYDELINSELCLPSLKHSGRELYLRTSSWAP 285
Qy 301 DKSRLRWEDEGNARLNLMLSYNMRGELAFLAENSDARGWEPLPERRLDAFRAIYNDWRGE 360
Db 301 DKSRLRWEDEGNARLNLMLSYNMRGELAFLAENSDARGWEPLPERRLDAFRAIYNDWRGE 342
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 4
US-07-894-212A-8
; Sequence 8, Application US/07894212A
; Patent No. 536883

```

; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-894-212A-8

Query Match      7.5%; Score 140.5; DB 1; Length 647;
Best Local Similarity 22.8%; Pred. No. 8.6e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYIVPIETLIKKEIP--FGLNITGYTLKFLP--KDIIIDVKGGIASDLIEIIGTSY 78
DB 27 EKCWMPLETL--BEYPMKVAIHTSGPLIEMLDONRPEYIDLRLSLVKGQVEIVVAGF 84
DB 79 THALPLPLSRVAQVORDREVE--ELFEVSPKGFVLPBLADPIIPALIKXNGYEYL 136
QY 137 FADGEAMLFSAHLNSAIPKIPLY-PHLIKAQRE-----KFRYISYLLGLRELARKA 187
DB 142 IVD-----DHFMSAGLSKEELWMPYTTDEGCVIAVFPIDELK--YIIPRPVDKV 192
QY 188 IKL--VFEKVTLKAV--KDIEAVPVVAVNTAVMLGIGRLPLMPKCVASWT----- 236
DB 193 LEYHSLIDGDESKVAVFHDGKEKFGIWPGETYEWY-----EKGWLRREFDR 239
QY 237 ---EDKNNILLYGDI-----FIGYRDIAGYRMSVBSGLLEVIDELNSELCLP----- 281
DB 240 ISSDEKINMLYTEYLEKYKRGVLVPIASV-----FEM-----SEWSLPAKQARLF 287
QY 282 ---SELKHSG---RELVLRTSSWAPDKSLRIWREDEGN---ARLNLSTYNNKGELAF 330
DB 288 VEFVNELKVGKIFPEKYVFPVARGIM---KNF-PKYKPSNVMHGMMLVSLVANN----- 339
QY 331 AENS DARGWEP LPERRLDAFPAIYND--WRG 359
DB 340 -----PEARKYLRLAQCNDAYWHG 358

```

```

RESULT 5
US-07-894-212A-2
; Sequence 2, Application US/07894212A
; Patent No. 5366883
; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MUKAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-894-212A-2

Query Match      7.5%; Score 140.5; DB 1; Length 649;
Best Local Similarity 22.8%; Pred. No. 8.7e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYIVPIETLIKKEIP--FGLNITGYTLKFLP--KDIIIDVKGGIASDLIEIIGTSY 78
DB 29 EKCWMPLETL--BEYPMKVAIHTSGPLIEMLDONRPEYIDLRLSLVKGQVEIVVAGF 86
QY 79 THALPLPLSRVAQVORDREVE--ELFEVSPKGFVLPBLADPIIPALIKXNGYEYL 136
DB 87 YEPVLASIP--KDRLEQIRLMKEMAKSIGFDARGVWLTERVQWPELVKTKSGGIDYV 143
QY 137 FADGEAMLFSAHLNSAIPKIPLY-PHLIKAQRE-----KFRYISYLLGLRELARKA 187
DB 144 IVD-----DHFMSAGLSKEELWMPYTTDEGCVIAVFPIDELK--YIIPRPVDKV 194
QY 188 IKL--VFEKVTLKAV--KDIEAVPVVAVNTAVMLGIGRLPLMPKCVASWT----- 236
DB 195 LEYHSLIDGDESKVAVFHDGKEKFGIWPGETYEWY-----EKGWLRREFDR 241
QY 237 ---EDKNNILLYGDI-----FIGYRDIAGYRMSVBSGLLEVIDELNSELCLP----- 281
DB 242 ISSDEKINMLYTEYLEKYKRGVLVPIASV-----FEM-----SEWSLPAKQARLF 289
QY 282 ---SELKHSG---RELVLRTSSWAPDKSLRIWREDEGN---ARLNLSTYNNKGELAF 330
DB 290 VEFVNELKVGKIFPEKYVFPVARGIM---KNF-PKYKPSNVMHGMMLVSLVANN----- 341

```

QY 331 AENSNDARGWEPLPERLDAFRAIYND--WRG 359
Db 342 -----PEARXYLLRAQCNDAYWVG 360

RESULT 6
US-07-893-928A-1
; Sequence 1, Application US/07893928A
; Patent No. 5578479
; GENERAL INFORMATION:
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: a-AMYLASE FROM HYPERTHERMOPHILIC
; TITLE OF INVENTION: ARCHABACTERIUM
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,928A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95470/C-1197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-893-928A-1

Query Match 7.5%; Score 140.5; DB 1; Length 650;
Best Local Similarity 22.8%; Pred. No. 8,7e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;
QY 25 EKAYIPVETLTKEIP---FGINIGYTLKLP---KDIIDLVKGASDLLEITIGTSY 78
Db 29 EKCWPFLELT--EETNMKVALHTSGPLELQDNRPETIDLRSLVKGQVEIVAGF 86
QY 79 THAILPLPLSRVQAQVDRREYKE--ELFEVSPKGFMLPELAYDPLIPAILKDNQYEVYL 136
Db 87 YEPVLTASIP--KEDRIEQIRLMKEMAKSIGFDARGWTLERWQGBELVYTKLESQIDYV 143
QY 137 FADGEANLFSAHLSAIPKIPKLY-PLHIAQRE-----KRFYISYLLGLRLRLRKA 187
Db 144 IVD-----DYHMSAGLSKEBLWYPTEDGGEVIAVFPIDKLR---YLIPFRVDRK 194
QY 188 IKL---VFEGKVTLKAV--KDIEAVPVWVAVNTAVNLGIGRLPLMNPKYASNT----- 236
Db 195 LEVLIHSLIDDESKAVAVFHDDGKFGIWPCTYEWV-----EKGGLREFFOR 241
QY 237 ---EDKDNILLYGTDI-----FIGYRDIAGYRMSVEGLLEVIDELINSELCLP----- 281
Db 242 ISSDEKINIMLYLEYLEKYPKGLVYLPASV-----FEM-----SEMSLPACQARLP 289
QY 282 ----SELKISG---RELVIKRTSSMAPDSLSLRIRWDEBN---ARINMLSYNMRGELAPL 330

Db 290 VEFVNELVKVGIFREKYRVFVRCGIM---KNF-FYKYPESVYMKRMILWYSKLVYRNN----- 341
QY 331 AENSNDARGWEPLPERLDAFRAIYND--WRG 359
Db 342 -----PEARXYLLRAQCNDAYWVG 360

RESULT 7
US-09-107-532A-5554
; Sequence 5554, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5554:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...227
; SEQUENCE DESCRIPTION: SEQ ID NO: 5554:
US-09-107-532A-5554

Query Match 5.6%; Score 105; DB 4; Length 227;
Best Local Similarity 24.2%; Pred. No. 0.0095;
Matches 60; Conservative 34; Mismatches 88; Indels 66; Gaps 12;
QY 140 GEAMLFSAHLSAIPKIPKLYPHLIIKAQREKRFYISYLLGLRLRLK-----AI 188
Db 3 GEKMIFSITMKLV-----FGLIGLLVVRLLGKKSSEITPPDLVY 43
QY 189 KLVEGKVTLKAVD-----IEAVPVWVAVNTAVNL-GIGRLPLMNPKYASWIEDKD 240
Db 44 TLVAGGILBESTYDNDVNVGHVLPALW-----AVMIYIGIERIVQKN-EKVNRWVGGE 97
QY 241 NILLYGTDIETFIGYRDIAGYRMSVEGLLEVIDELINSELCLPSE-----LKHSGRLELYR 294

Db 98 SVLIKD---GVINNTLNTNNHIEMSQLRAI---LRQCECFLENAKHVILENAGQMSVLK 151
Qy 295 TSSWAPDKSLIFMREDEGNALNTLSYMRGELAFIANSDARGHEPRLDLAFRIY 354
Db 152 KSD--EDPAUSILLVDEGOIQHKVLSQNLTE-AMLMENLKKEGY-----ADVKOLIT 201
Qy 355 NDWRGENG 362
Db 202 VEWSEKRG 209

RESULT 8

US-09-710-279-2150
Sequence 2150, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2150
LENGTH: 360
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2150

Query Match 5.4%; Score 101.5; DB 4; Length 360;
Best Local Similarity 19.1%; Pred. No. 0.047;
Matches 57; Conservative 52; Mismatches 117; Indels 73; Gaps 12;

Qy 4 LVFHGNQVAFEPKSEIPKVIKAYIVETLKEEIPFGNTGYTLKPLPKDIIDLK 63
Db 61 LIFRG-----VKRIVEDGYSIIRKLIONNI--NLIALHTNLDVNPKGVNRLA 107
Qy 64 GGASDLIEIGT--SYTHAIIPLPLSRVBAQVQRDEVEBELFEVSPKFWMLPELAYD 121
Db 108 DQGLENTSMINTSSYYKVQTFIPKVIIE-----DFKDSINEL----- 147
Qy 122 PIIPAILKNDGYEYLFADGEMLFSAHLNSAIKPKLYPHLIRAKQEKRRFRYSYLLGL 181
Db 148 ---GLAKEGNYEYCFPESSEG-----KQGFKEVGDASFYIGKLD-----IEYV--- 187
Qy 182 RELRKAIKLVF---EGKVTLKAVKDIE--AVPV--WVAVNTAVMLGIGRLPLMPKCV 232
Db 188 ---DEIKLEFMKIDNELITKRAILDNHPEYTFVPDFPIKKNKSEVGLGIIGQLNQTM 243
Qy 233 ASWIEDKNILLYGTDFIEFGYRDIAGYRMSVEGLELV-----IDELNSELCLPSELKH 286
Db 244 LDFSEYAKQKQNLIPSVRYTGQHDSPIKVAIIIGSGGIFPEYKASQIGADVFTGDIKH 302

RESULT 9

US-09-710-279-2106
Sequence 2106, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2106
LENGTH: 376
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2106

Query Match 5.4%; Score 101.5; DB 4; Length 376;
Best Local Similarity 19.1%; Pred. No. 0.05;
Matches 57; Conservative 52; Mismatches 117; Indels 73; Gaps 12;

Qy 4 LVFHGNQVAFEPKSEIPKVIKAYIVETLKEEIPFGNTGYTLKPLPKDIIDLK 63
Db 47 LIFRG-----VKRIVEDGYSIIRKLIONNI--NLIALHTNLDVNPKGVNRLA 93
Qy 64 GGASDLIEIGT--SYTHAIIPLPLSRVBAQVQRDEVEBELFEVSPKFWMLPELAYD 121
Db 94 DQGLENTSMINTSSYYKVQTFIPKVIIE-----DFKDSINEL----- 133
Qy 122 PIIPAILKNDGYEYLFADGEMLFSAHLNSAIKPKLYPHLIRAKQEKRRFRYSYLLGL 181
Db 134 ---GLAKEGNYEYCFPESSEG-----KQGFKEVGDASFYIGKLD-----IEYV--- 173
Qy 182 RELRKAIKLVF---EGKVTLKAVKDIE--AVPV--WVAVNTAVMLGIGRLPLMPKCV 232
Db 174 ---DEIKLEFMKIDNELITKRAILDNHPEYTFVPDFPIKKNKSEVGLGIIGQLNQTM 229
Qy 233 ASWIEDKNILLYGTDFIEFGYRDIAGYRMSVEGLELV-----IDELNSELCLPSELKH 286
Db 230 LDFSEYAKQKQNLIPSVRYTGQHDSPIKVAIIIGSGGIFPEYKASQIGADVFTGDIKH 288

RESULT 10

US-08-270-013B-2
Sequence 2, Application US/08270013B
Patent No. 5686294
GENERAL INFORMATION:
APPLICANT: Sogabe et al.
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Volt & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 61601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/270,013B
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 164701/1993
FILING DATE: 02-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Robert F.
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 62321
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: (25)3533
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-270-013B-2

Query Match 5.4%; Score 101; DB 1; Length 329;
Best Local Similarity 25.3%; Pred. No. 0.046;
Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;

QY 10 LQYAEIPKSEIPK-----VIEKAYIPVETLKEIPFGINITYTLKFLPKDII DL 61
DB 130 MTYVFKESGFPKRVIGSGVLDTR---FRFVAEE---LNIS-----VKDYVGF 175
QY 62 VKGIASDLIEIGTSTHAI--LPLPLSRVDAQVQREVEKELEFVSPKG--FWLPE 117
DB 176 VLGGHGDWVPLVRSYAGIPIPLEKLI PKRDIAVERTRKGGGEIVNLGNSAYYAPA 235
QY 118 LAYDPIIPALIKDN-----GYEYLF 137
DB 236 ASLVEMVEALIKDQRRILPAIAYLEGEYEGIV 269

RESULT 11

US-08-838-418-2
Sequence 2, Application US/08838418
Patent No. 5744342
GENERAL INFORMATION:
APPLICANT: Sogabe et al.
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,418
FILING DATE: 17-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,013
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 164701/1993
FILING DATE: 02-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Allen E.
REGISTRATION NUMBER: 37354
REFERENCE/DOCKET NUMBER: 78339
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: (25)3533
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-838-418-2

Query Match 5.4%; Score 101; DB 1; Length 329;
Best Local Similarity 25.3%; Pred. No. 0.046;
Matches 39; Conservative 25; Mismatches 50; Indels 40; Gaps 7;

QY 10 LQYAEIPKSEIPK-----VIEKAYIPVETLKEIPFGINITYTLKFLPKDII DL 61
DB 130 MTYVFKESGFPKRVIGSGVLDTR---FRFVAEE---LNIS-----VKDYVGF 175
QY 62 VKGIASDLIEIGTSTHAI--LPLPLSRVDAQVQREVEKELEFVSPKG--FWLPE 117
DB 176 VLGGHGDWVPLVRSYAGIPIPLEKLI PKRDIAVERTRKGGGEIVNLGNSAYYAPA 235
QY 118 LAYDPIIPALIKDN-----GYEYLF 137
DB 236 ASLVEMVEALIKDQRRILPAIAYLEGEYEGIV 269

RESULT 12

US-09-540-236-3128
Sequence 3128, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA.
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 3128
LENGTH: 653
TYPE: PR
ORGANISM: M. catarrhalis
US-09-540-236-3128

Query Match 5.2%; Score 97; DB 4; Length 653;
Best Local Similarity 21.9%; Pred. No. 0.37;
Matches 72; Conservative 43; Mismatches 128; Indels 86; Gaps 13;

QY 22 KYIEKAYIPVETLKEIPFGINITYTLKFLP-----KDIIDLVKGIAS--- 68
DB 95 KRIEPAHYGVITD--REGAPLAANPVTVPFAVAEYRYLDNEIKTKSETAKOKA 152
QY 69 -----DLIEIGTSTHAIPLPLSRVDAQVQD-----REVEKELEFVSPKG 113
DB 153 LKKLKEMDLVR-----AAANYPLEKLEAAVIGIDHTLDTNSQKKEAL----PKGA 201
QY 114 WLPELA-YDPIIPALKNGIEYLFADGEAMLPSAHLNSAIKPIPLYPHLIKAKREKRF 172
DB 202 SSRRLVTLNRRVSPVAKSVTDLGLFALIGREOYFORYYQA--EPNAQLIGVAAQSSDFTQG 260
QY 173 RYISYLLGLRELKRAIKVFGSKVTLKAVKDI EAVPVVAANTVMTGIGLPLMPKRV 232
DB 261 GYIG-----RAGIEAKYERLAGKDKG-----VOILKGTGQPIQELIQI 300
QY 233 ASWIEDKONILLYGTDIEFIFYRDIAGYRMSVEGLEVIDELNSLCLPSLEKHSGRLEY 292
DB 301 EPLIEGENIRLITDSRLQYVLYKE-----LEQVRLQ-----SARSSSGMWVD 343
QY 293 LRT-----SSMAPDKSIRIRBEGNAR 315
DB 344 VKTGEVTLAMGSPSPFNNSNLSERDGNAR 372

RESULT 13

US-09-134-001C-3678
Sequence 3678, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: CTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3678
 ; LENGTH: 367
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3678

Query Match 5.0%; Score 94.5; DB 3; Length 367;
 Best Local Similarity 18.7%; Pred. No. 0.27;
 Matches 56; Conservative 52; Mismatches 118; Indels 73; Gaps 12;

QY 4 LVFIGNLOYAIPKSEIPKVIKAYIPYIEFLIKEIPFGNITTYTLKPLKDIIDLVK 63
 DB 68 LIFKG-----VRIVEDGGSIIIRKLIQNNI--NLIALHTNIDVNPKGNNMLA 114
 QY 64 GIASDLIEIGT--SYTHAILPLPLSRVEAQQVORDREVKELPEVSPKGFMLPELAYD 121
 DB 115 DQIGENISMINTSSYYKQTPFKNYIE-----DFKDSLNEL----- 154
 QY 122 PIPAILKONGEYELFADGEMLSAHLNSAIKPIKPLYPHLIKAQREKFRYISYLLGL 181
 DB 155 ---GLAKGNYECFFESG-----KGGFPGVGDASPTGKLD-----IEYV--- 194
 QY 182 RELKAIKLVF---EKKVTLKAVKDE--AVPV--WVAVNTAVMLGIGRLPLMPKKV 232
 DB 195 ---DEIDLEFMIKONLEITRKAILDHPYETPVPFDIKKKNKESEYGLGIIQOLNQTMT 250
 QY 233 ASWIEDKONILLYGTDFIGYRDIAGYRMSVEGLEV-----IDELNSELCLPSELKH 286
 DB 251 LDSESEVAKKQNLNIPSVARYTGOHDSPIKVAIIIGSGIGFHYKASQLGADVFTGDIKH 309

RESULT 14

US-09-710-279-1840
 ; Sequence 1840; Application US/09710279
 ; Patent No. 6703492
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUS480US
 ; CURRENT APPLICATION NUMBER: US/09/710,279
 ; PRIOR FILING DATE: 2000-11-09
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1840
 ; LENGTH: 375
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 US-09-710-279-1840

Query Match 4.9%; Score 92; DB 4; Length 375;
 Best Local Similarity 19.3%; Pred. No. 0.51;
 Matches 72; Conservative 55; Mismatches 109; Indels 138; Gaps 16;

QY 59 IDLVKGIASDLIEIGTSYTHAILPLPLSRV---EAQQVORDREVKELPEVSPK--- 111
 DB 3 IDISSG---DFIAFIGTSGSGKTTALMINRMIESTGEITIDGKNIKEINLPVELRSI 58
 QY 112 GFWEPELAYD-----IIPAILK-----DNGEYV----- 136
 DB 59 GYVIOQGLMPHMTVKENIVVPPKLLKWSQEKKEKAKELIRLVLPPEEYIDRYPSSELG 118
 QY 137 ---FADGEMLSAHLNSAIKPI--KPLYPHLIK---AQREKFRYISYLLGL 181
 DB 119 GQQRIGVVRALAAEQDIIILMDEPFGALDPIITRDTLDLVKQLQQLAKTIFVTH--- 174

QY 182 RELKAIKLT-----VFEGKVTLKAVKD 203
 DB 175 -DMDEAIKLDKICIMTNGQVITYDTPDNILRSFANDFVDFIQGNRIQRPNIRYKD 233
 QY 204 IEAVPVVAVNTAVMLGIGRLPLNPKVASMIBDKONILLYGTDFIEFGRIAGYRMS 263
 DB 234 AMIRPVYVHDSRLANDAVN---IMREKV-----DTIPVGNDEHLGLDIEDINEG 283
 QY 264 VEGLEVIDELNSELCL---LPSELKHSGRELYLTSSWAP-----DKSL----- 304
 DB 284 LRHHKELIDTQORDIYRVRIKSDSVRTIKGNVRNVPVSDNKTLLGLVTRANLV 343
 QY 305 ---RIVRE--DEGN 313
 DB 344 IVDYSIWGELESNG 357

RESULT 15

US-09-248-796A-15394
 ; Sequence 15394; Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Kelch Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBIC
 ; FILE REFERENCE: 107136.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 15394
 ; LENGTH: 845
 ; TYPE: PRT
 ; ORGANISM: Candida albicans
 US-09-248-796A-15394

Query Match 4.9%; Score 92; DB 4; Length 845;
 Best Local Similarity 20.2%; Pred. No. 1.9;
 Matches 78; Conservative 64; Mismatches 117; Indels 128; Gaps 21;

QY 38 EIRP-FGLNITGYTLKFLPPDIIDLKVGIASDLIEITL-----GTSYTHAI 82
 DB 202 DEMPHIGVDING-----KRIMPAKGSALDQLLESIDLPEGWTGLDQWGTGS----- 249
 QY 83 LPLPLSRVEAQQVORDREVKELPE--VSPKGFMLPELAYDPIIPAILKONGEYELFADGE 141
 DB 250 ---LKTDBELELELRKIQOQENTDENINP-----YELIDWPTKD----- 286
 QY 142 AMLPSAHLNSAIKPIKPLYPHLIKAQREKFRYISYLLGRELKAIKLVPEGVK----- 196
 DB 287 -----EIRMP-----VNAVPEPKRRFPSPGHEAKRWIKYAIRREGIIPPNK 329
 QY 197 ---TKAVKDIKAVPVM---VAVNTAVM--LGIGRLP-----LMPKKVA 233
 DB 330 VKQQLTSEEBEDQNFPLMDOEIRISDHIMWLRAPKLPPTNEESVNPPEEYLLTSEKS 389
 QY 234 SWIE---DKDNILLYGTDFIEFGRIAGYRMSVEGLEVIDELNSELCLPSELKISGR 289
 DB 390 KWLQESPDIRRNFPL---PKYNSLRQVPGYQDSVRRF---ERSIDLVLAPVRHN-- 440
 QY 290 ELVYRTSSWAPD-----KSLR-----IWRDEGNARLMLNLSYNNKGELAFLAENG-- 335
 DB 441 KLNIDPSLPLDLPSPDLAPFPPIRCSTIYEGHTG--KIRTISIDPQG--LMLATGSDDG 496
 QY 336 -ARGMEPLPERRLDAPRAIYNMKGEN 361
 DB 497 SVRIWEILTGROYVYKQILNKKEINNEED 523

Search completed: January 27, 2005, 06:00:12
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 28, 2005, 04:41:08 ; Search time 573 Seconds
(without alignments)
3650.093 Million cell updates/sec

Title: US-09-886-400A-4
Perfect score: 1879
Sequence: 1 LRALVFHGNQYVAIPKSEI.....RLDAPRAYNDWNGENGP 364

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool.p/US09886400/runat_27012005_055724_17979/app_query.fasta_1.519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09886400@CGN_1_1_723@runat_27012005_055724_17979
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq: *
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq: *
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------------------|--------------------|
| 1 | 1879 | 100.0 | 1095 | 15 US-10-112-231A-3 | Sequence 3, Appl1 |
| 2 | 1870 | 99.5 | 1095 | 9 US-09-886-400-3 | Sequence 3, Appl1 |
| 3 | 1870 | 99.5 | 1095 | 13 US-10-112-357-3 | Sequence 3, Appl1 |
| 4 | 1870 | 99.5 | 1095 | 13 US-10-114-403-3 | Sequence 3, Appl1 |
| 5 | 1870 | 99.5 | 1095 | 13 US-10-116-606-3 | Sequence 3, Appl1 |
| 6 | 1870 | 99.5 | 1095 | 13 US-10-112-331-3 | Sequence 3, Appl1 |
| 7 | 1870 | 99.5 | 1095 | 13 US-10-112-377-3 | Sequence 3, Appl1 |
| 8 | 1870 | 99.5 | 1095 | 13 US-10-116-581-3 | Sequence 3, Appl1 |
| 9 | 1870 | 99.5 | 1095 | 13 US-10-112-442-3 | Sequence 3, Appl1 |
| 10 | 1870 | 99.5 | 1095 | 13 US-10-112-418-3 | Sequence 3, Appl1 |
| 11 | 1870 | 99.5 | 1095 | 13 US-10-114-083-3 | Sequence 3, Appl1 |
| 12 | 111.5 | 5.9 | 2670 | 16 US-10-282-122A-17097 | Sequence 17097, A |
| 13 | 109 | 5.8 | 939 | 15 US-10-369-493-46924 | Sequence 46924, A |
| 14 | 108 | 5.7 | 2455 | 16 US-10-425-114-29659 | Sequence 29659, A |
| 15 | 108 | 5.7 | 2586 | 16 US-10-424-599-32675 | Sequence 32675, A |
| 16 | 106 | 5.6 | 1770 | 18 US-10-425-115-106717 | Sequence 106717, A |
| 17 | 105 | 5.6 | 822 | 9 US-09-974-300-56 | Sequence 56, Appl1 |
| 18 | 103 | 5.5 | 3222 | 9 US-08-815-242-8263 | Sequence 8263, Ap |
| 19 | 103 | 5.5 | 3441 | 9 US-09-815-242-4413 | Sequence 4413, Ap |
| 20 | 103 | 5.5 | 3453 | 16 US-10-282-122A-8207 | Sequence 8207, Ap |
| 21 | 103 | 5.5 | 5030 | 8 US-08-781-986A-324 | Sequence 324, App |
| 22 | 103 | 5.5 | 5030 | 16 US-10-329-624-324 | Sequence 324, App |
| 23 | 102.5 | 5.5 | 1434 | 15 US-10-369-493-28193 | Sequence 28193, A |
| 24 | 102.5 | 5.5 | 1434 | 15 US-10-369-493-30951 | Sequence 30951, A |
| 25 | 102.5 | 5.5 | 3036 | 16 US-10-398-221-2003 | Sequence 2003, Ap |
| 26 | 101 | 5.4 | 1326 | 17 US-10-437-963-20238 | Sequence 20238, A |
| 27 | 101 | 5.4 | 4755 | 16 US-10-398-221-3741 | Sequence 3741, Ap |
| 28 | 100.5 | 5.3 | 945 | 15 US-10-369-493-41082 | Sequence 41082, A |
| 29 | 100 | 5.3 | 1345 | 16 US-10-425-114-32190 | Sequence 32190, A |
| 30 | 100 | 5.3 | 1441 | 18 US-10-425-115-136289 | Sequence 136289, A |
| 31 | 99.5 | 5.3 | 994 | 16 US-10-425-114-5764 | Sequence 5764, Ap |
| 32 | 99.5 | 5.3 | 994 | 16 US-10-425-114-5766 | Sequence 5766, Ap |
| 33 | 99 | 5.3 | 536165 | 10 US-09-939-964-1 | Sequence 1, Appl1 |
| 34 | 98.5 | 5.2 | 1800 | 15 US-10-369-493-45612 | Sequence 45612, A |
| 35 | 98.5 | 5.2 | 1800 | 18 US-10-793-639-209 | Sequence 209, App |
| 36 | 98.5 | 5.2 | 2442 | 16 US-10-282-122A-11024 | Sequence 11024, A |
| 37 | 98.5 | 5.2 | 3082 | 13 US-10-113-852A-1 | Sequence 1, Appl1 |
| 38 | 98 | 5.2 | 1533 | 15 US-10-369-493-23743 | Sequence 23743, A |
| 39 | 97.5 | 5.2 | 1342 | 18 US-10-425-115-51870 | Sequence 51870, A |
| 40 | 97.5 | 5.2 | 1590 | 15 US-10-369-493-33955 | Sequence 33955, A |
| 41 | 97 | 5.2 | 1955 | 16 US-10-282-122A-26871 | Sequence 26871, A |
| 42 | 97 | 5.2 | 2925 | 15 US-10-094-749-581 | Sequence 581, App |
| 43 | 97 | 5.2 | 100848 | 16 US-10-672-787-39 | Sequence 39, Appl1 |
| 44 | 96.5 | 5.1 | 2498 | 16 US-10-425-114-20468 | Sequence 20468, A |
| 45 | 96.5 | 5.1 | 2914 | 18 US-10-739-930-1249 | Sequence 1249, Ap |

ALIGNMENTS

RESULT 1
US-10-112-231A-3
; Sequence 3, Application US/10112231A
; Publication No. US20030184276A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/112,231A
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 09/407,806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 06/613,220
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA

ORGANISM: Thermococcus alcaliphilus
US-10-112-231A-3

Alignment Scores:

| Pred. No.: | 2.27e-209 | Length: | 1095 |
|------------------------|-----------|---------------|------|
| Score: | 1879.00 | Matches: | 364 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 15 | Gaps: | 0 |

US-09-886-400A-4 (1-364) x US-10-112-231A-3 (1-1095)

```

QY      1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db      1 TTGAGAGGGCTCGCTTTCACGGCAACCTCCAGTATGCCAAATCCCAAGAGCGGAATC 60

QY      21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db      61 CCAAGGTCATAGAGAGCATACATCCAGTCCAGACACTGATTAAGAAATTT 120

QY      41 ProheGlyLeuAniIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db      121 CTTTGGGCTCAACATACGAGGCTATACCTTAAAGTCTCCCGAAGATATTATAGAC 180

QY      61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db      181 CTGGTTAAAGGGGCATCGGAGAGTACCTGATAGATATCCGAAGACCTACCGCAC 240

QY      81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
Db      241 GCATATCTCCCGCTCCCGCTTACGAGTTCAGAGTCAAGACCAAGTTCAGAGATAGGAA 300

QY      101 ValLysGluGluLeuPheGluValSerProLysGlyPheThrLeuProGluLeuAlaTyr 120
Db      301 GTTAAAGAAAGCTCTTCGAGGTTTCTCCAAAGGATTCGCGTCCGACGCTCGCTAT 360

QY      121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
Db      361 GACCCGATATCTCCGTCATCTGAGAGACACACGGTTATGATATCTATTCGCGCACGG 420

QY      141 GluAlaMetLeuPheSerAlaHisLeuAniSerAlaIleLysProIleLysProLeuTyr 160
Db      421 GAGCGATGCTTTTCTCAGCTCATCTCACTCGCGCATTAAGCCAAATTAAACGCTCTAT 480

QY      161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
Db      481 CCACACCTTATTAAGCCCAAGGAAAGCGCTTTAGGATCACTACGCTATCTCTTGGT 540

QY      181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
Db      541 CTCAGGAGGCTTACGAGAGCGATTAAGCTGTTTGAAGGTAAAGCTTAACGCTTAAGCA 600

QY      201 ValLysAspIleGluAlaValProValITPrValAlaValAsnThrAlaValMetLeuGly 220
Db      601 GTCAAAAGACATCAAAACCGTACCCGTTGGGTGGCCGTAAACACGCTGTATATGCTCGCG 660

QY      221 IleGlyArgLeuProLeuMetAsnProLysValAlaSerTrpIleGluAspLysAsp 240
Db      661 ATCGGAAGGCTTCTCTTATGATCTTAAGAAAGTGGCAGCTGGATAGAGCAAGAGAC 720

QY      241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
Db      721 AACATCTCTTATACCGCACCGCATATAGAGTTCATGCTATAGGACATTCGAGGCTAC 780

QY      261 ArgMetSerValGluLysLeuGluValIleAspGluLeuAniSerGluLeuCysLeu 280
Db      781 AGATATAGGTTGAGGATATTATAGGTTATAGAGACTCAACCTCGGACCTGTGCTT 840

QY      281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300
Db      841 CCTCAGAGCTGAGACAGTGAAGGAGCTCTACTTACGACCTTCGAGTTGGGACCA 900
  
```

```

QY      301 AspLysSerLeuArgIleTrpArgGluAspGluIleAsnAlaArgLeuAniMetLeuSer 320
Db      901 GATTAAGGCTTGGATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

QY      321 TyrAsnMetArgGlyGlyLeuLeuAlaPheLeuAlaGluAniSerAspAlaArgGlyTrpGlu 340
Db      961 TACAAATATAGAGGCGCAACTCCGCTTTTATGCCAGAACAGCGATGCAAGGAGATGGAG 1020

QY      341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
Db      1021 CCCCTCCCTGAGAGAGAGGCTGATGCTTCCGGGCGATATATACGATTAAGAGGAGTGA 1080

QY      361 AsnGlyGluPro 364
Db      1081 AATGGGAACT 1092
  
```

RESULT 2

```

US-09-886-400-3
; Sequence 3, Application US/09886400
; Patent No. US20020045226A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1092)
US-09-886-400-3
  
```

Alignment Scores:

| Pred. No.: | 2.56e-208 | Length: | 1095 |
|------------------------|------------------------------------------|---------------|------|
| Score: | 1870.00 <td>Matches:</td> <td>362</td> | Matches: | 362 |
| Percent Similarity: | 99.73% <td>Conservative:</td> <td>1</td> | Conservative: | 1 |
| Best Local Similarity: | 99.45% <td>Mismatches:</td> <td>0</td> | Mismatches: | 0 |
| Query Match: | 99.52% <td>Indels:</td> <td>0</td> | Indels: | 0 |
| DB: | 9 <td>Gaps:</td> <td>0</td> | Gaps: | 0 |

US-09-886-400A-4 (1-364) x US-09-886-400-3 (1-1095)

```

QY      1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db      1 TTGAGAGGGCTCGCTTTCACGGCAACCTCCAGTATGCCAAATCCCAAGAGCGGAATC 60

QY      21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db      61 CCAAGGTCATAGAGAGCATACATCCAGTCCAGACACTGATTAAGAAATTT 120

QY      41 ProheGlyLeuAniIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db      121 CTTTGGGCTCAACATACGAGGCTATACCTTAAAGTCTCCCGAAGATATTATAGAC 180

QY      61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db      181 CTGGTTAAAGGGGCATCGGAGAGTACCTGATAGATATCCGAAGACGATACACGCAC 240

QY      81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
  
```


QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
 Db 721 AACATCTCTTACGACGACGATATAGATTTCATGGCTATAGGACATTGCAAGGCTAC 780
 QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
 Db 781 AGAATGAGTGTGAGGAGATTATTAGAGGTTATAGCGACTCACTCGGAACTGTGCTT 840
 QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300
 Db 841 CCTCAGAGCTGAAGACAGTGAAGGAGGACTTACCTTACCGACTTCGAGTTGGGCACCA 900
 QY 301 AspLysSerLeuArgIleTyrArgGluAspGluGluAsnAlaArgLeuAsnMetLeuSer 320
 Db 901 GATAGAGCTTGAGGATATGAGAGAGACGAAAGGAAACGCAACCTTAATATCTGTCC 960
 QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
 Db 961 TACAAATATGAGAGGCGCAACTGCCCTTTTACCGAAGAACGATGCAAGGCGATGGAG 1020
 QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360
 Db 1021 CCCCCTCCCTGAGAGAGGCTGGATGCCCTTCGCGGCGATATATACGATTGGAGGGGTGAA 1080
 QY 361 AsnGlyGluPro 364
 Db 1081 AATGGGGAACCT 1092

RESULT 4

US-10-114-403-3
 ; Sequence 3, Application US/10114403
 ; Publication No. US20020115100A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Ried, John
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
 ; FILE REFERENCE: DIVER1120-4
 ; CURRENT APPLICATION NUMBER: US/10/114,403
 ; PRIOR FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: 09/886,400
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 09/619,032
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/407,806
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 08/613,220
 ; PRIOR FILING DATE: 1996-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1095
 ; TYPE: DNA
 ; ORGANISM: Thermococcus alcaliphilus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1092)
 ; US-10-114-403-3

Alignment Scores:

Pred. No.: 2,566-208 Length: 1095
 Score: 1870.00 Matches: 362
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 99.52% Indels: 0
 DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-114-403-3 (1-1095)

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1081 AATGGGGAACCT 1092

Db 1 TTGAGACCGCTGCTTTTACAGGCAACCTCCAGTATCCGAATCCCAAGACGGAATC 60
 QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluIle 40
 Db 61 CCAAGGCTCATAGAGAGGATATCATCCAGTCATGAGACACTGATTAAAGAAATTT 120
 QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
 Db 121 CTTTTGGGCTCAACATTAACGGGCTATACCTTAAAGTTCTCCCGAAGATATTATAGAC 180
 QY 61 LeuValLysGlyGlyIleIleAsnSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
 Db 181 CTGCTTAAGGGGGCATTCGGAATGCTGATAGCATTAATGGAACGAGCTACACGAC 240
 QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluValGlnArgAspArgGlu 100
 Db 241 GCATATCTCCCTCTCCCGCTTAGCAGATAGAAAGCAAGATTGAGAGATAGGAA 300
 QY 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTyrLeuProGluLeuAlaTyr 120
 Db 301 GTTAAAGAAAGCTCTTCGAGCTTCTCCAAAGGGAATTCGCTGCCAGAGCTCGCTAT 360
 QY 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
 Db 361 GACCCGATTAATCCCTGCCATCTGAAAGACAAAGCTTATGATATCTATTCGCCGACGG 420
 QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160
 Db 421 GAGCGATGCTTTTCTCAGCTCATCTCAACTCGGGGATTAAGCCATTAACCGCTCAT 480
 QY 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLysGly 180
 Db 481 CCACACTTATTAAGGCCCAAGAGGAAAGCCCTTTAGGTATCATCAGCTATCTCTTGAT 540
 QY 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
 Db 541 CTCAGGAGCTTAGAGAGCGCATTAAGCTCGTTTGAAGGTAAAGCTTAAGGCA 600
 QY 201 ValLysAspIleGluAlaValProValIleTyrValAlaValAsnThrAlaValMetLeuGly 220
 Db 601 GTCAAGACATCGAAGCCGTAACCCGTTGGGGCGCGGAACAACGCTGTAATGCTCGGC 660
 QY 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspLysAsp 240
 Db 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGAGCTGATGAGGACAAAGAC 720
 QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
 Db 721 AACATCTCTTATACGGACCGATATAGATTCACTTGGGACATTGGAGCATTCGAGCTAC 780
 QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
 Db 781 AGAATGAGTGTGAGGAGATTATTAGAGGTTATAGCGACTCACTCGAAGCTGCGCTT 840
 QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300
 Db 841 CCTCAGAGCTGAAGACAGTGAAGGAGGACTTACCTTACCGACTTCGAGTTGGGCACCA 900
 QY 301 AspLysSerLeuArgIleTyrArgGluAspGluGluAsnAlaArgLeuAsnMetLeuSer 320
 Db 901 GATAGAGCTTGAGGATATGAGAGAGACGAAAGGAAACGCAACCTTAATATCTGTCC 960
 QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
 Db 961 TACAAATATGAGAGGCGCAACTGCCCTTTTACCGAAGAACGATGCAAGGCGATGGAG 1020
 QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360
 Db 1021 CCCCCTCCCTGAGAGAGGCTGGATGCCCTTCGCGGCGATATATACGATTGGAGGGGTGAA 1080
 QY 361 AsnGlyGluPro 364
 Db 1081 AATGGGGAACCT 1092

RESULT 5

US-10-116-606-3
 ; Sequence 3, Application US/10116606
 ; Publication No. US20020119515A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Ried, John
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
 ; FILE REFERENCE: DIVER1120-4
 ; CURRENT APPLICATION NUMBER: US/10/116,606
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US/09/886,400
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 09/619,032
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/407,806
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 08/613,220
 ; PRIOR FILING DATE: 1996-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1095
 ; TYPE: DNA
 ; ORGANISM: Thermococcus alcaliphilus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1092)
 ; US-10-116-606-3

Alignment Scores:

Pred. No.: 2,566-208 Length: 1095
 Score: 1870.00 Matches: 362
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 99.52% Indels: 0
 DB: Gaps: 0

US-09-886-400a-4 (1-364) x US-10-116-606-3 (1-1095)

Qy 1 LeuArGAlaLeuValPheHISglYAsnLeuGlnTYrAlaGluIleProLysSerGluIle 20
 Db 1 TTGAGAGCGCTGCTGCTTTTCACGCAACTCCAGTATGCCAAATCCCAAGGCGAAATTC 60
 Qy 21 ProLysValIleGluValAlaTYrIleProValIleGluThrIleLysGluGluIle 40
 Db 61 CCAAGGTCATAGAGAGGCAATACATCCAGTCATCGAGCACTGATTAAAGAAATTT 120
 Qy 41 ProPheGlyLeuAsnIleThrGlyTYrThrLeuLysPheLeuProLysAspIleIleAsp 60
 Db 121 CCTTTGGGCTCAACATTAACGGGCTATACCTTAAAGTTCCCTCCGAAAGATATTATAGAC 180
 Qy 61 LeuValLysGluValIleAlaSerAspLeuIleGluIleIleGlyThrSerTYrThrHis 80
 Db 181 CTCGTAAAGGGGGGATCGCAAGTACCTGATAGGAATTCGAGACAGCTACCGCAC 240
 Qy 81 AlaIleLeuProLeuLeuProLysSerArgValGluAlaGlnValAlaArgAspArgGlu 100
 Db 241 GCATATACCTCCCTCTCTCCGCTTAGCAGATAGAGCAAGATTCAAGAGATAGGGA 300
 Qy 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTYr 120
 Db 301 GTTAAAGGAAGAGCTTCTGAGCTTTCTCCAAAGGGAATTCGGCTGCCAGAGCTCGCTAT 360
 Qy 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTYrGluTYrTrpLeuPheAlaAspGly 140
 Db 361 GACCCGATTAATCCCTGCACTACTGAAAGCAACGGTATGATATCTATTTGCGCAGCGG 420
 Qy 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTYr 160
 Db 421 GAGGGAGAGCTTTTCTCAGCTCATCTCACTCGGGAATTAAGCCAATTAAACCGCTCAT 480

Qy 161 ProHisLeuIleLysAlaGlnArgGluLysAspPheArgTYrIleSerTYrLeuLeuGly 180
 Db 481 CCACACTTATTAAGGCCCAAGAGGAGGCTTTAGTACATCAGCTATCTCTGTGT 540
 Qy 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
 Db 541 CTCAGGAGCTTAGGAAGCGCATTAAGCTGTTTTTGAAGGTAAAGTAAACCTAAAGCA 600
 Qy 201 ValLysAspIleGluAlaValProValTrpValAlaValAsnThrAlaValMetLeuGly 220
 Db 601 GTCAAGACATCGAAGCGGTACCCGTTGGGTGGCCCGTGAACACGGCTGTAATCTCGGC 660
 Qy 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTrpIleGluAspLysAsp 240
 Db 661 ATCGAAGGCTTCTCTTATGAATCTTAAGAAAGTGGGAGCTGATAGAGCAAGGAC 720
 Qy 241 AsnIleLeuLeuTYrGlyThrAspIleGluPheIleGlyTYrArgAspIleAlaGlyTYr 260
 Db 721 AACATTCTTATACGGCAGCATATAGATTGCTATAGGACATTGCAAGGCTAC 780
 Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
 Db 781 AGAATGAGGTGTGAGGAGTATTAGAGGTATTAGACGAGCTCACTCGAACTGTGCTT 840
 Qy 281 ProSerGluLeuLysHisSerGlyArgGluLeuTYrLeuArgTYrThrSerTrpAlaPro 300
 Db 841 CCTCAGAGCTGAAGCAGAGTGAAGGAGGCTTCACTTACGAGCTCAGTTGGGCACCA 900
 Qy 301 AspLysSerLeuArgIleTrpArgLysAspGluLysAsnAlaArgLeuAsnMetLeuSer 320
 Db 901 GATTAAGACTTGAAGATATGAGAGAGAGCAAGAGCAAGCACTTAATCTGTCTGCC 960
 Qy 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTrpGlu 340
 Db 961 TACATATGAGGGGCGAAGCTGCCCTTTTACCGCAAGCAAGCATGCAAGGAGTGGAG 1020
 Qy 341 ProLeuProGluAlaGluLeuAspAlaPheArgAlaIleTYrAsnAspTrpArgGlyGlu 360
 Db 1021 CCCCTCCCTGAGAGAGGCTGATGCTTCGGGCGATATTAACGATTTGAGGGGTGAA 1080
 Qy 361 AsnGlyGluPro 364
 Db 1081 AATGGGAACTT 1092

RESULT 6
 ; US-10-112-331-3
 ; Sequence 3, Application US/10112331
 ; Publication No. US20020119550A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Ried, John
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE 1
 ; FILE REFERENCE: DIVER1120-4
 ; CURRENT APPLICATION NUMBER: US/10/112,331
 ; PRIOR FILING DATE: 2002-03-29
 ; PRIOR APPLICATION NUMBER: US/09/886,400
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 09/619,032
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/407,806
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 08/613,220
 ; PRIOR FILING DATE: 1996-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1095
 ; TYPE: DNA
 ; ORGANISM: Thermococcus alcaliphilus
 ; FEATURE:
 ; NAME/KEY: CDS

```

; LOCATION: (1)...(1092)
US-10-112-331-3

Alignment Scores:
Pred. No.: 2,56e-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-112-331-3 (1-1095)

QY 1 LeuAArgAlaLeuValPheHisGlyAseNLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db 1 TTGAAGAGCGCTCGCTCTTTCACGGCAACCTCCAGATAGCCGAATCCCAAGAGCGGAATC 60

QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db 61 CCAAGGCTCATAGAGAGGCGATACATCCAGTCATCGACACACTGATTAAGAAATTT 120

QY 41 ProPheGlyLeuAseNileThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db 121 CTTTTGGGCTCAACATTAACGGGCTATACCTTAAGATTCTCCCGAAGGATATTATAGAC 180

QY 61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db 181 CTGGTTAAAGGGGCGATCGCGAGTACCTGATAGATTAATCGAAGCGACTACAGCAC 240

QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnatGAspArgGlu 100
Db 241 GCATATACCTCCCTCCCTCCCGCTTAGAGAGTAGAAGCAAAATTCAGAGATAGAGGA 300

QY 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTyr 120
Db 301 GTTAAAGAGAGGCTCTTCGAGCTTTCTCCAAAGGGATTCGGCTGCCAGAGCTCGCCTAT 360

QY 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGlnTyrLeuPheAlaAspGly 140
Db 361 GACCCGATTAATCCCTCGCATACGAAAGAACACGGTATAGATTCATTCGCGGAGGG 420

QY 141 GluAlaMetLeuPheSerAlaHisLeuAseNserAlaIleLysProIleLysProLeuTyr 160
Db 421 GAGGCGATGCTTTTCTCAGCTCANTCTCACTCGGCGAATAAGCAATTAACGGCTTAT 480

QY 161 ProHisLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
Db 481 CCACACCTTAAAGGCCCAAGGGAAGCGCTTTAGTACATCACTATCTCTTGCT 540

QY 181 LeuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
Db 541 CTGAGGAGCTTAGAGAGCGAATTAACCTGCTTTTAAGAGTAAAGGTAACGCTTAAGGCA 600

QY 201 ValLysAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
Db 601 GTCAAAACATTCGAAGCGCTACCGCTTTGGGTGGCCGTGAACCGCTGTAATGCTCGC 660

QY 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTrpIleGluAspLysAsp 240
Db 661 ATCGGAAGGCTTCTCTTATGAATCTTAAGAAAGTGGAGCTGGAATAGAGGACAAAGAC 720

QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
Db 721 AACATTTCTTAAGGCGACCGAATTAAGATTCATTTGCTATAGGAGCATTTGCGAGCTAC 780

QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAseNserGluLeuCyLeu 280
Db 781 AGATAGATGTGTGAGGATTAATTAAGGTTATTAACAGAGCTCAACTCGGAATGTGCTT 840

QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300
Db 841 CCTCAAGAGCTGAAGCACAGTGAAGGAGAGCTTACTTACGAGCTTGAAGTTGGGACCA 900

```

```

QY 301 AspLysSerLeuArgIleTrpArgLysAspGluIleAsnAlaArgLeuAseNMetLeuSer 320
Db 901 GATTAAGAGCTTAGAGATATAGAGAGCGAAGGGAACGCAAGACTTATATATGCTGTCC 960

QY 321 TyrAseNMetArgGlyGluLeuAlaPheLeuAlaGluAseNserAspAlaArgGlyTyrGlu 340
Db 961 TCAATATATAGAGGCGGCAACTCCCTTTTAGCCGAGAACAGGAGTGCAGAGGGGATGGAG 1020

QY 341 ProLeuProGluLysArgLysAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
Db 1021 CCCCTCCCTGAAGAGGAGGCTGATGCTTCCGGGCGATATATACGATTTGAGGGGTGA 1080

QY 361 AsnGlyGluPro 364
Db 1081 AATGGGGAACCT 1092

RESULT 7
US-10-112-377-3
; Sequence 3, Application US/10112377
; Publication No. US20020120108A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,377
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1092)
US-10-112-377-3

Alignment Scores:
Pred. No.: 2,56e-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-112-377-3 (1-1095)

QY 1 LeuAArgAlaLeuValPheHisGlyAseNLeuGlnTyrAlaGluIleProLysSerGluIle 20
Db 1 TTGAAGAGCGCTCGCTCTTTCACGGCAACCTCCAGATAGCCGAATCCCAAGAGCGGAATC 60

QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
Db 61 CCAAGGCTCATAGAGAGGCGATACATCCAGTCATCGACACACTGATTAAGAAATTT 120

QY 41 ProPheGlyLeuAseNileThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
Db 121 CTTTTGGGCTCAACATTAACGGGCTATACCTTAAGATTCTCCCGAAGGATATTATAGAC 180

QY 61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
Db 181 CTGGTTAAAGGGGCGATCGCGAGTACCTGATAGATTAATCGAAGCGACTACAGCAC 240

```



```

Qy 81 AlaileuProleuLeuProleuSerArgValGluValGlnValGlnArgAspArgGlu 100
Db 241 GCAATACCTCCCTCTCTCCGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 101 VallysgluGluLeuPheGluValSerProlysglyPheTrpLeuProGluLeuValTyr 120
Db 301 GTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 121 AspProilelleProAlaileuLeuValSerAspArgGlyTyrGluTyrLeuPheAlaAspGly 140
Db 361 GACCGAGATATCCCTGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 141 GluAlaMetLeuPheSerAlaileuLeuValSerProlysglyPheTrpLeuProGluLeuValTyr 160
Db 421 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 161 ProHileuileuValGlnArgGluValSerArgPheArgTyrIleSerTyrLeuLeuGly 180
Db 481 CCACACCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 181 LeuArgGluLeuValGlnArgGluValSerArgPheArgTyrIleSerTyrLeuLeuGly 200
Db 541 CTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 201 VallyAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
Db 601 GTCAAGACATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 221 IleGlyArgLeuProleuMetAsnProlysglyValAlaSerTrpIleGluAspArg 240
Db 661 ATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
Db 721 AACATCTCTTATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
Db 781 AGAATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 281 ProSerGluLeuValSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300
Db 841 CCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 301 AspLySerLeuArgIleTyrArgGluAspGlyValAsnAlaArgLeuAsnMetLeuSer 320
Db 901 GATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
Db 961 TACAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 341 ProleuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
Db 1021 CCCCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 361 AsnGlyGluPro 364
Db 1081 AATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

```

; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1092)
US-10-116-581-3

Alignment Scores:
Pred. No.: 2,566-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-116-581-3 (1-1095)
Qy 1 LeuArgAlaLeuValPheIleGlyAsnLeuGlnTyrAlaGluIleProlySerGluLe 20
Db 1 TTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 21 ProlyValIleGluValAlaTyrIleProValIleGluThrLeuIleGluGluIle 40
Db 61 CCAAGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 41 ProHeglyLeuAsnIleThrGlyTyrThrLeuValSerPheLeuProlyAspIleLeu 60
Db 121 CTTTGGGCTCAACATTAACGGGCTTATCTTAAAGTTCCTCCGAGAGATATATGAC 180
Qy 61 LeuVallysglyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThr 80
Db 181 CTCGTTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 81 AlaileuProleuLeuProleuSerArgValGluValGlnValGlnArgAspArgGlu 100
Db 241 GCAATACCTCCCTCTCTCCGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 101 VallysgluGluLeuPheGluValSerProlysglyPheTrpLeuProGluLeuValTyr 120
Db 301 GTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 121 AspProilelleProAlaileuLeuValSerAspArgGlyTyrGluTyrLeuPheAlaAspGly 140
Db 361 GACCGAGATATCCCTGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 141 GluAlaMetLeuPheSerAlaileuLeuValSerProlysglyPheTrpLeuProGluLeuValTyr 160
Db 421 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 161 ProHileuileuValGlnArgGluValSerArgPheArgTyrIleSerTyrLeuLeuGly 180
Db 481 CCACACCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 181 LeuArgGluLeuValGlnArgGluValSerArgPheArgTyrIleSerTyrLeuLeuGly 200
Db 541 CTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 201 VallyAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
Db 601 GTCAAGACATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

```

QY 221 ILeGlyAArgLeuProLeuMetAsnProLySylsValAlaSerTriPleGluAspLysAsp 240
Db 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGGAGCTGATAGAGCAAGGAC 720
QY 241 AsnIleLeuLeuTyrgLyThraSpIleGluPheIleGlyTyraAspIleAlaGlyTyr 260
Db 721 AACATTTCTTCAATCGGACCCGATATAGATTCAATTGGCTATAGGACATTCAGGCTAC 780
QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
Db 781 AGAATGAGTGTGGAGGATTTATAGAGTTATAGACAGCTCACTGGAACTGTGTCTT 840
QY 281 ProSerGluLeuLySylSerGlyAArgGlyLeuTyrgLeuAArgThraSerSerTriPAlaPro 300
Db 841 CCTCAGAGCTGAAGCAAGTGAAGGAGCTCACTTACGAGCTTGAGTGGGACCA 900
QY 301 AspLySerLeuAArgIleTyrgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320
Db 901 GATAGAGCTTGAAGATATGAGAGAGACGACAGGAAAGCAAGACTTAATATGTCTTCC 960
QY 321 TyrAsnMetAArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
Db 961 TACATATGAGGGGCGAACTGCGCTTTAGCCGAGAAACAGCTGCAAGGGATGGAG 1020
QY 341 ProLeuProGluAArgAArgLeuAspAlaPheAArgAlaIleTyraAsnAspTriPArgGlyGlu 360
Db 1021 CCCCTCCCTGAGAGAGAGCTGATGCTTCCGGGCGATATATTAACGATTGGAGGGGTGA 1080
QY 361 AsnGlyGluPro 364
Db 1081 AATGGGAACTT 1092

RESULT 9

US-10-112-442-3

Sequence 3, Application US/10112442

Publication No. US20020150997A1

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: Murphy, Dennis

APPLICANT: Ried, John

TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH

CURRENT FILING DATE: 2002-03-29

CURRENT FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: US/10/112,442

PRIOR FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: 09/886,400

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/619,032

PRIOR FILING DATE: 1999-09-20

PRIOR APPLICATION NUMBER: 08/613,220

PRIOR FILING DATE: 1996-03-08

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1095

TYPE: DNA

ORGANISM: Thermococcus alcaliphilus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1092)

US-10-112-442-3

Alignment Scores:

Pred. No.: 2,566-208

Score: 1870.00

Percent Similarity: 99.73%

Best Local Similarity: 99.45%

Query Match: 99.52%

DB: 13

Length: 1095

Matches: 362

Conservative: 1

Mismatches: 1

Indels: 0

Gaps: 0

US-09-886-400A-4 (1-364) x US-10-112-442-3 (1-1095)

QY 1 LeuAArgAlaLeuValPheHisGlyAenLeuGluIleTyrAlaGluIleProLySerGluIle 20
Db 1 TTGAGAGCGCTCGTCTTTACCGGCAACTCCAGTATGCGGAATCCCAAGAGCGAAATC 60
QY 21 ProLyValIleGluLySylAArgIleProValIleGluThreLeuIleGlyGluIle 40
Db 61 CCAAGGCTATAGAGAGCAATACATCCAGATCATGAGACACTGATTAAAGAAATTT 120
QY 41 ProPheGlyLeuAsnIleThrgLyTyThraLeuLyPheLeuProLySylIleAsp 60
Db 121 CCTTTGGGCTCAACATTAACGGGCTATACCTTAAGTCTTCCCGAAGATATTATAGAC 180
QY 61 LeuValLyGlyGlyIleAlaSerAspLeuIleGluIleIleGlyTyThraSerTyThraHis 80
Db 181 CTGTTAAAGGGGCAATCGAGTGCCTGATAGATATATCGAAACGAGCTACAGGCAC 240
QY 81 AlaIleLeuProLeuLeuProLeuSerAArgValGluAlaGluValGlnAArgAspArgGlu 100
Db 241 GCAATATCTCCCTCTCCGCTTACAGATAGAAAGCAAGTTCAGAGATAGAGAA 300
QY 101 ValLyGlyGluLeuPheGluValSerProLyGlyPheTyrgLeuProGluLeuAlaTyrg 120
Db 301 GTTAAAGGAAGCTCTTCGAGCTTCTCCAAAGGATTCGTGGCTGCAGAGCTCGCTAT 360
QY 121 AspProIleIleProAlaIleLeuLySylAspAsnGlyTyrgGluTyrgLeuPheAlaAspGly 140
Db 361 GACCCGATATATCCCTGCGCATCTGAAAGCAACGTTAGATATCTATTCGCCACGGG 420
QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLyProIleLySylProLeuTyrg 160
Db 421 GAGGCGATGCTTTTCTGAGCTCATCTCAACTGGGATTAACCAATTAAACGCTCAT 480
QY 161 ProHisLeuIleLyAlaGluAArgGlyTyrgAArgPheAArgTyrgIleSerTyrgLeuGly 180
Db 481 CCACACCTTATTAAGGCCCAAGGAGAAAGCCCTTATGATCATCATGATCTCTTGGT 540
QY 181 LeuAArgGluLeuAArgLySylAlaIleLySylLeuValPheGlyGlyValThreLeuLySyl 200
Db 541 CTCAGGAGCTTATAGAAAGCGATTAAGCTCGTTTGAAGGATTAACGCTTAAGGGA 600
QY 201 ValLyAspIleGluAlaValProValTyrgAlaValAsnThraAlaValMetLeuGly 220
Db 601 GTCAAAAGCATGAGAGCGGTACCGCTTGGGGGCGGTGAACAGCGCTGATATGCTCGG 660
QY 221 ILeGlyAArgLeuProLeuMetAsnProLySylsValAlaSerTriPleGluAspLysAsp 240
Db 661 ATCGAAGGCTTCTCTTATGATCTTAAGAAAGTGGGAGCTGATAGAGCAAGGAC 720
QY 241 AsnIleLeuLeuTyrgLyThraSpIleGluPheIleGlyTyraAspIleAlaGlyTyr 260
Db 721 AACATTTCTTCAATCGGACCCGATATAGATTCAATTGGCTATAGGACATTCAGGCTAC 780
QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
Db 781 AGAATGAGTGTGAAGATTAATTAGAGTTATAGAGCTTAAGAGCTCAACTCGGAAGTGTCTT 840
QY 281 ProSerGluLeuLySylSerGlyAArgGlyLeuTyrgLeuAArgThraSerSerTriPAlaPro 300
Db 841 CCTCAGAGCTGAAGCAAGTGAAGGAGCTCACTTACGAGCTTGAGTGGGACCA 900
QY 301 AspLySerLeuAArgIleTyrgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320
Db 901 GATAGAGCTTGAAGATATGAGAGAGACGACAGGAAAGCAAGACTTAATATGTCTTCC 960
QY 321 TyrAsnMetAArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
Db 961 TACATATGAGGGGCGAACTGCGCTTTAGCCGAGAAACAGCTGCAAGGGATGGAG 1020
QY 341 ProLeuProGluAArgAArgLeuAspAlaPheAArgAlaIleTyraAsnAspTriPArgGlyGlu 360
Db 1021 CCCCTCCCTGAGAGAGAGCTGATGCTTCCGGGCGATATATTAACGATTGGAGGGGTGA 1080
QY 361 AsnGlyGluPro 364

Db 1081 AATGGGAACCT 1092

RESULT 10

US-10-112-418-3
 ; Sequence 3, Application US/10112418
 ; Publication No. US20020155486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Ried, John
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
 ; FILE REFERENCE: DIVER1120-4
 ; CURRENT APPLICATION NUMBER: US/10/112,418
 ; CURRENT FILING DATE: 2002-03-29
 ; PRIOR APPLICATION NUMBER: 09/886,400
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 09/619,032
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/407,806
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 08/613,220
 ; PRIOR FILING DATE: 1996-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1095
 ; TYPE: DNA
 ; ORGANISM: Thermococcus alkaliphilus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1092)
 ; US-10-112-418-3

Alignment Scores:

Pred. No.: 2,566-208 Length: 1095
 Score: 1870.00 Matches: 362
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 99.52% Indels: 0
 Gaps: 0

US-09-886-400a-4 (1-364) x US-10-112-418-3 (1-1095)

Qy 1 LeuAArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluIleProLysSerGluIle 20
 Db 1 TTGAGAGCGCTCGCTTTCACGGCAACCTCCAGTATGCCGAATCCCAAGAGCGGAATC 60
 Qy 21 ProLysValIleGluValAlaTyrIleProValIleGluThrLeuIleLysGluGluIle 40
 Db 61 CCAAGGTCATAGAGAGGCAATACCTCCAGTATCCGAGACATGATTAAGAGAAATT 120
 Qy 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleIleAsp 60
 Db 121 CTTTGGGCTCAACATTAACGGGCTATACCTTAAGTTCCCTCCGGAAGGATTTATAGAC 180
 Qy 61 LeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
 Db 181 CTCCTTAAGGGGCAATCGCAGTACCTGATAGATTAATCGGAACGAGCTACACGCAC 240
 Qy 81 AlaIleLeuProLeuLeuProLeuSerArgValGluIleValGluIleValGluAspArgGlu 100
 Db 241 GCATTAATCTCCCTCCCTCCCTCCGCTTAAGCAGATTAAGAGCAAGTTCCAGAGATAGGAA 300
 Qy 101 ValLysGluGluLeuPheGluValSerProLysGlyPheTyrPleuProGluLeuAlaTyr 120
 Db 301 GTTAAGAGAGGCTCTTCGAGCTTTCTCCAAAGGATTTCTGGCTGCCAGAGCTGCCCTAT 360
 Qy 121 AspProIleIlePheProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
 Db 361 GACCCGATTAATCCCTCGCATCTAGTGAAGAGCAACGTTATGATATCTATTCGCCAGCGGG 420

Qy 141 GluAlaMetLeuPheSerAlaIleLeuAsnSerAlaIleLysProIleLysProLeuTyr 160
 Db 421 GAGCGAGATGCTTTTCTAGCTCATCTCACTCGCGCATTAAGCCATTAAGCCGCTCAT 480
 Qy 161 ProHISLeuIleLysAlaGlnArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
 Db 481 CCACACCTTAATAAGGCCCAAGGAAAGGCGTTAGTACATCAGCTATCTCCTTGGT 540
 Qy 181 LeuArgGluLeuAspGlyValIleLysLeuValPheGluGlyLysValThrLeuValAla 200
 Db 541 CTCAGGAGCTTAGAAGGCGATTAAGCTCTTTTGAAGGTAAAGTAAACCTTAAGGCA 600
 Qy 201 ValLysAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
 Db 601 GTCAAGAGCATCGAAGCCGTACCCGTTGGGTGGCCGTGAACAGCGCTGTAATCTCGGC 660
 Qy 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspLysAsp 240
 Db 661 ATCGAAGGCTTCTCTTATGATATCTTAAGAAAGTGGCGAGCTGATAGACACAAAGAC 720
 Qy 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
 Db 721 AACATTCCTTATACGGCACCGATATAGATTCAATTGGCTATAGGCAATTGCAAGCTAC 780
 Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
 Db 781 AGATGAGTGTGGAGGATTTTGAAGTTTAAACAGCTCACTCGAATCTGCTT 840
 Qy 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTyrAlaPro 300
 Db 841 CCTCAGAGCTGAAGCACAGTGAAGGAGGAGCTTACTTACGAGCTTGAGTGGGACCA 900
 Qy 301 AspLysSerLeuAlaGlyIleTyrArgGluAspGluLysAlaArgLeuAsnMetLeuSer 320
 Db 901 GATTAAGCTTGAAGATATGAGAGAGACCAAGGAAACCGCAACTTAATATCTGTGCC 960
 Qy 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
 Db 961 TACATATGAGGAGGCGAATCTGCCCTTTTACCCAGAAACCGATGCAAGGAGATGGGAG 1020
 Qy 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArgGlyGlu 360
 Db 1021 CCCCTCCCTGAGAGAGGCTGAGATGCCCTTCGGGCGATATATTAAGATTGGAGGGTGCA 1080
 Qy 361 AsnGlyGluPro 364
 Db 1081 AATGGGAACCT 1092

RESULT 11

US-10-114-083-3
 ; Sequence 3, Application US/10114083
 ; Publication No. US20020160464A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Ried, John
 ; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
 ; FILE REFERENCE: DIVER1120-4
 ; CURRENT APPLICATION NUMBER: US/10/114,083
 ; CURRENT FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: 09/886,400
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: 09/619,032
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/407,806
 ; PRIOR FILING DATE: 1999-09-20
 ; PRIOR APPLICATION NUMBER: 08/613,220
 ; PRIOR FILING DATE: 1996-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1095

```

; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1092)
US-10-114-083-3

```

```

Alignment Scores:
Pred. No.: 2,566-208 Length: 1095
Score: 1870.00 Matches: 362
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

```

US-09-886-400a-4 (1-364) x US-10-114-083-3 (1-1095)

```

QY 1 LeuArgAlaLeuValPheHisGlyAsnLeuGlnTyrAlaGluLeuProLysSerGluLeu 20
DB 1 TTGAGAGCGCTCTCTTTTCACGGCACTTCCAGTATGCCAAATCCCAAGAGCGCAATC 60
QY 21 ProLysValIleGluLysAlaTyrIleProValIleGluThrLeuIleLysGluLeu 40
DB 61 CCAAGGTCATAGGAAGGATACATCCAGTCATCGACACTGATTAAAGAAAT 120
QY 41 ProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuProLysAspIleLeu 60
DB 121 CTTTGGGCTCAACATACGAGCTATACCTTAAAGTCTCCCGAAGGATATTATAGAC 180
QY 61 LeuValLysGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHis 80
DB 181 CTGTTAAAGGGGCGATCGCGAGTACTGATAGAGTATCCGAAGAGACTACCGCAC 240
QY 81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluValGluArgAspArgGlu 100
DB 241 GCATACTCCCTCCCTCCCGCTTACGAGTAGAACACAAAGTTACAGAGATAGGGA 300
QY 101 ValLysGluLeuLeuPheGluValSerProLysGlyPheThrLeuProGluLeuAlaTyr 120
DB 301 GTTAAAGGAAGGCTTTCGAGCTTTCCTCAAGGAGATTTCGGCTCCAGAGCTCGCTAT 360
QY 121 AspProIleIleProAlaIleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGly 140
DB 361 GACCCGATATCCCTCGCATACGAAAGACACGCTTATGATATCATTCGCGACGG 420
QY 141 GluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyr 160
DB 421 GAGCGATGCTTTCTCAGCTCATCTCAACTCGCGATTAAGCCAAATTAAACCGCTCAT 480
QY 161 ProHisLeuIleLysAlaGluArgGluLysArgPheArgTyrIleSerTyrLeuLeuGly 180
DB 481 CCACACCTTAAAGGCCCAAGGAAAGGCGCTTTCAGTACATCACTTCTCTTGGT 540
QY 181 LeuArgLysLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrLeuLysAla 200
DB 541 CTCGAGGAGCTTACGAAGCGATTAAGCTCTTTTGAAGGTAAAGTAAAGCAAGCA 600
QY 201 ValLysAspIleGluAlaValProValTyrValAlaValAsnThrAlaValMetLeuGly 220
DB 601 GTCAAAACATCGAAGCGTACCGCTTGGGTGCGCCGTGAACACGGCTGTATCTCGGC 660
QY 221 IleGlyArgLeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspAsp 240
DB 661 ATCGAAGGCTTCTCTTATATATCTTAAGAAAGTGGAGCTGATAGAGAAAGGAC 720
QY 241 AsnIleLeuLeuTyrGlyThrAspIleGluPheIleGlyTyrArgAspIleAlaGlyTyr 260
DB 721 AACATTCTTCTATACGCAACCGATATAGATTCACTTGGCTATAGGACATTCGAGGCTAC 780
QY 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeu 280
DB 781 AGAATGAGGTGTTGAGGAGTATATAGAGGTATATAGCGAGCTCAACTCGGAATGTGCTT 840

```

```

QY 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaPro 300
DB 841 CCTTAGAGCTGACACACAGTGAAGGAGCTTACTTACGAGCTTCAGATTGGCACCA 900
QY 301 AspLysSerLeuArgIleTrpArgLysAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320
DB 901 GATTAAGAGCTTGAAGATATGAGAGAGACGAAAGCAAGCACTTAATATGTGTGCC 960
QY 321 TyrAsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGlu 340
DB 961 TACCAATATGAGGGGCGAACTCCCTTTAGCGGAACAGGATGCAAGGAGATGGAG 1020
QY 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTrpArgGlyGlu 360
DB 1021 CCCCTCCCTGAGAGAGAGGTGGATGCCCTTCGCGGCGATATATACATTGAGGGGTGA 1080
QY 361 AsnGlyGluPro 364
DB 1081 AATGGGGAACCT 1092

```

RESULT 12

```

US-10-282-122A-17097
; Sequence 17097, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELIIRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17097
; LENGTH: 2670
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-282-122A-17097

```

```

Alignment Scores:
Pred. No.: 0.0446 Length: 2670
Score: 111.50 Matches: 72

```

| | | | |
|------------------------|-------|---------------|----|
| Percent Similarity: | 39.3% | Conservative: | 48 |
| Best Local Similarity: | 23.6% | Mismatches: | 98 |
| Query Match: | 5.93% | Indels: | 87 |
| DB: | 16 | Gaps: | 16 |

US-09-886-400A-4 (1-364) X US-10-282-122A-17097 (1-2670)

| | | | |
|----|------|-----------------------------------------------------------------|--------|
| QY | 53 | PheLeuProLysAspIleIleAspLeuValIysGlyIleIleAspLeuIleu | 72 |
| Db | 1664 | TTCCACACA---GATTATATAATAGCTGTAGGTGAGAGAGTCATGATGATCTGGTAA | 17404 |
| QY | 73 | IleIleGlyThrSerTyThrHisAlaIleuProLeuProLeuSerArgValIu | 92 |
| Db | 1741 | ATAATGTGGGTAAATGTATGACATCCAGAGATGATTTCCAAAGCTTAGCT----- | 179191 |
| QY | 93 | AlaGlnValGlnArgAspArgGluValIysGluGluLeuPheGluValSerProIysGly | 112 |
| Db | 1792 | -----ATGAGATTATATGGATATTAAGAAAAGAGATAT--GATATTCCTAAAG-- | 18368 |
| QY | 113 | PheThrLeuProGluLeuAlaTyArgProIleIlePro----- | 125 |
| Db | 1837 | -----ATGGGTGAAAAGCGGATTTTGTCTGTATCCCACTTCAGCAGAACTGGTTCA | 18996 |
| QY | 126 | -----AlaIleLeu-----LysAspAsnGlyTyArgIuTyLeuPheAla | 138 |
| Db | 1891 | GAAGTAATCCATTTGGCAGTTATACTGACCAAGATTCAGGTGTAATACCATCCATTGCA | 19506 |
| QY | 139 | AspGlyGluAlaMetLeuPheSerAlaHisLeuAsnSerAlaIleTyProIleArgPro | 158 |
| Db | 1951 | GATATATGAATTAATGCCCAACATGGCTTAATATGATGCTGATATGATATGAGATGCCA | 201010 |
| QY | 159 | LeuTyProHisIleu-----IleLysAlaGlnArgGluLysArgPheArg | 173 |
| Db | 2011 | -----CCAAATTAACAGCAGCATCAGGTGTGATGCTTTAAACATGATTCAGAAACA | 20644 |
| QY | 174 | TyrIleSerTyLeu-----LeuGlyLeuArgGluLeuArgLys | 186 |
| Db | 2065 | TATGATATTAATGTTAAGAACTGAACACGCTGACGTCTTACGACTA-----CAA | 21122 |
| QY | 187 | AlaIleLysLeuValPheGlu-----GlyLysValThrLeu | 198 |
| Db | 2113 | GCAGGTAAATAATATTTGAAATATTACCTGTGCTATTAATAATGAAAGAAATGATAA | 21727 |
| QY | 199 | LysAlaValLysAspIleGluAlaValProValTyrValAlaValAsnThrAla---- | 216 |
| Db | 2173 | GAAACCAAGAAAAAATGSCAAATGGCATCTACTATGCGAAGAAATGCTTTTGCAAAATGCC | 22322 |
| QY | 217 | -----ValMetLeu | 219 |
| Db | 2233 | TTCTTAGGAAATGTGCACTCTTTAGCACATAAATTAGTGCAATTCACCATGTTCAACAC | 22922 |
| QY | 220 | GlyIleGlyArgLeuProLeuMetAsn-----Pro | 229 |
| Db | 2293 | GGTGTGCAAAATGCAATTTATTAATAATGAAGTAATCAAGTTCAATTTGCGAAGAGCTCA | 23524 |
| QY | 230 | LysIleValAlaSerTyrIleGluAspLys-----AspAsnIleLeuLeuTyGlyThr | 247 |
| Db | 2353 | AATAAATAGGGTGCACTTCTCACAATACAGATATCTGTGATTCATACAAAGATATGCT-- | 24099 |
| QY | 248 | AspIleGluPheIleGlyTyArgAspIleAlaGlyTy-----ArgMetSerValGlu | 265 |
| Db | 2410 | -----GATTTGCTCTTTTTCGAGCGATTTAAAGGAATACAGACCAAGAAAAGTATGAT | 24633 |
| QY | 266 | GlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuTybLeuProSerGluLeuLys | 285 |
| Db | 2464 | AATCTTAATAAAGCAATATGATGAATTAAGCGCAAAAGTATGGCTTTCTTAAGACTATATAAA | 25233 |
| QY | 286 | HisSerGlyArgGlu | 290 |
| Db | 2524 | GAGGCTGAGTAA | 2538 |

```

Sequence 46924: Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 46924
LENGTH: 939
TYPE: DNA
ORGANISM: Bacillus subtilis
US-10-369-493-46924

```

US-09-886-400A-4 (1-364) X US-10-369-493-46924 (1-939)

| | | | | |
|----|-----|---------------------------------------------------------------|-------|-----|
| Oy | 10 | LeuGlnTyrAlaGluIleProIysSerGluIleProIys | ----- | 22 |
| Oy | | | | |
| Db | 388 | ATGACATACGCGCGGTACAAAGAAATCAGGCTTCCCTTAAGAGCGGTAAATCGGCCACTCA | ----- | 447 |
| Oy | 23 | ---ValIleGluLysAlaTyrIleProValIleGluThrIleuLeuYsgIuGluIlePro | 41 | |
| Db | 448 | GGGTGCTCTGATACGGCAAGA-----TTCAAGACATTGTGGCAGAGAA----- | 492 | |
| Oy | 42 | PheGlyLeuAsnIleThrGlyTyrThrIleuLysPheLeuProLysAspIleIleAspLeu | 61 | |
| Db | 493 | -----TTAAACCTGTCA-----GTGAAGAATGTGACTGCTTTC | 525 | |
| Oy | 62 | ValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGlyThrSerTyrThrHisAla | 81 | |
| Db | 526 | GTHCTGGCGGACACGGGTACGATATGGTTCCTGCGCTTATCTTATGCTGGCGGT | 585 | |
| Oy | 82 | Ile-----LeuProLeuLeuProLeuSerArgValGluIleGlnValGlnArgAspArg | 99 | |
| Db | 586 | ATCCGCGTTGAACCTTATTCTCGAAGAACGAGATGGATCGCAATGTGGAGCGGCACTAGA | 645 | |
| Oy | 100 | GluValLysGluGluLeuPheGluValSerProLysGly-----PheThrLeuProGlu | 117 | |
| Db | 646 | AAAGGCGAGCGGAATCGTGAATCTTCTTGGAAACGGAGCGCGATTTATGCGCGTCCG | 705 | |
| Oy | 118 | LeuAlaTyrAspProIleIleProAlaIleLeuLysAspAsn | 131 | |
| Db | 706 | GCTTCTGTGACAGAAATGTCGAGACGATCTTGGAAAGATCAGCGCGCTTCTCTTCA | 765 | |
| Oy | 132 | -----GlyTyrGluTyrLeuPhe | 137 | |
| Db | 766 | ATTGCTTATCTGAAGGGGAATACGCGTATATAGAAGCATCTACCTGTGTCTTCTACAAAT | 825 | |
| Oy | 138 | -----AlaAspGlyGluAlaMetLeuPheSer | 147 | |
| Db | 826 | GTAAGCGCGAAGGGCTTGTAGCAAAATCATTTAACTGAACCTGACAGACTATGAAGAAGCG | 885 | |
| Oy | 148 | HisLeuAsnSerAlaIleLysProIleLys | 157 | |
| Db | 886 | CAGCTGAATTAATTCAGTTGAATCTGTCAA | 915 | |

```

US-10-425-114-29659
; Sequence 29659, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Yindong
; APPLICANT: Zhou, Yinhua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ. ID NOS: 73128
; SEQ ID NO 29659
; LENGTH: 2455
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GWFLMINSOY109F02_FLI
US-10-425-114-29659

Alignment Scores:
Pred. No.:      0.101      Length:      2455
Score:          108.00     Matches:      85
Percent Similarity: 36.73%   Conservative: 59
Best Local Similarity: 21.68% Mismatches: 106
Query Match:      5.75%     Indels:      142
DB:               16       Gaps:       21

US-09-886-400a-4 (1-364) x US-10-425-114-29659 (1-2455)

QY      6 PheHisGlyAsnLeu---GlnTyrAlaGluIleProLysSer-----Glu 19
DB      1415 TTTCTGGGATCAGCTTGGAAGCTGTGGTTCATCAACCTCGGGGCTGAGTGGGAG 1474
QY      20 IleProLysValIleGluLysAlaTyrIleProVal----- 31
DB      1475 GTCCCTCCCAATGTGACCCCACTTTTCTCCATATTATTGTTGGTCCCGCAGGATT 1534
QY      32 -----IleGluThrLeu 35
DB      1535 TCTACCTGCTCCTTCATGAGCAGATTACCGCAATCCCAACATCATATTGAACACTG 1594
QY      36 IleLysGluGluIleProPheGlyLeuAsnIleThrGlyTyrThrLeuLysPheLeuPro 55
DB      1595 -----CCTTCCCTCAGGTCCTCTTATTATGACCAATCCTCATCG----- 1636
QY      56 LysAspIleIleAspLeuValLysGlyGlyIleAlaSerAspLeuIleGluIleIleGly 75
DB      1637 -----GTAACTTGTG-----GATCCAGTACAACCTTCGGT 1669
QY      76 ThisSerTyrThrHisAlaIleLeuProLeuLeuProLeuSerArgValGluAlaGluVal 95
DB      1670 ACTTCATGAGCACTTCTCT-----CCTTTTTCCT-----GATCCAGTGGTC 1711
QY      96 GlnArgAspArgGluValLysGluGluLeuPheGluValSerProLysGlyPheTrpLeu 115
DB      1712 AGAATGCTCTTGGCAACATCTCAGACATTCACCAATTCACACCA----- 1756
QY      116 ProGluLeuAlaLysrProIle-----IleProAlaIle---LeuLysAspAsnGly 132
DB      1757 -----TTAATGTGTATCCAAATGTTTCATGTCCAGTGTACGTGCTCTTCAGGT 1810
QY      133 TyrGluTyrLeuPheAlaAspGlyGluAlaMetLeuPheSerAlaHisLeuAsnSerAla 152
DB      1811 CAGGGCTATCTTGTAGAGTGTGGCCCTGCAATGCACCTTGC----- 1852
QY      153 IleLysProIleLysProLeuTyrProHisLeuIleLysAlaGlnArgGluLysArgPhe 172
DB      1853 -----ATTCCACCATGTGATCCAAACCTTGTGAGGACATGATTCTGTGAATCTGAT 1903

```

```

QY      173 ArgTyrIleSerTyrLeuLeuGlyLeuArgGluLeuArgLysAlaIleLysLeuValPhe 192
DB      1904 GCTGTGGTG-----AAAGGTGCAAGAGAGACTTGAATTGCTTACTT 1945
QY      193 GluGly-----LysValThrLeuLysAlaValLysAsp 203
DB      1946 AGTGGTTCAAGCAGGAGGTAAACAGAGATGATAGGAGACCTGCTGCATTT----- 1999
QY      204 IleGluAlaValProValIleTyrValAlaValAsnThrAlaValMetLeuGlyIleArg 223
DB      1999 ----- 1999
QY      224 LeuProLeuMetAsnProLysLysValAlaIleSerTrpIleGluAspLysAspAsnIleLeu 243
DB      2000 -----TTAAGTAAATTCAGAC-----GAAATTCAAAACATATTCCTT 2035
QY      244 LeuTyrGlyTyrAspIleIleGluPheIleGlyTyrArgAspIle----- 257
DB      2036 GTGGCTGGAAGTGTGTCTTACACTGACAGACAGATATTATCTTATTCGAAACAGC 2095
QY      258 -----AlaGlyTyrArgSerValGluGlyLeuLeuGluValIleAspGluLeuAsn 275
DB      2096 ATGTGCTGATAGGATTAAGTATGATGCTGTCTGGGGGTATCCAGGTAGACAGTGAAGTCTAT 2155
QY      276 SerGluLeuGlyLeuProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgThr 295
DB      2156 TCGGACTATATGT-----GAAACTATGAGGAATTTGGACAGCAGGAAATAAT 2200
QY      296 SerSerTrpAlaProAspLysSerLeuArgIleTyrArgGluAspGluGlyAsnAlaArg 315
DB      2201 TCCAAC-----GATTCAGGTGTGTGCTTTTGTGATGATGAGGCGGATCTTCT 2251
QY      316 Ieu-----AsnMetLeuSerTyrAsnMetArgGly 325
DB      2252 TTGAGATTGAGAGAGTGAAGTGAATGAATGCAATGTCTTTTCTTTTATTATGTA 2311
QY      326 GluLeuAlaPheLeuAlaGluAsnSerAspAlaArg 337
DB      2312 GATCTCTTTTGTGTAGCA---AATGTATGATATGCCG 2344

RESULT 15
US-10-424-599-32675
; Sequence 32675, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ. ID NOS: 285684
; SEQ ID NO 32675
; LENGTH: 2586
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12950C.1
US-10-424-599-32675

Alignment Scores:
Pred. No.:      0.109      Length:      2586
Score:          108.00     Matches:      85
Percent Similarity: 36.73%   Conservative: 59
Best Local Similarity: 21.68% Mismatches: 106
Query Match:      5.75%     Indels:      142
DB:               16       Gaps:       21

US-09-886-400a-4 (1-364) x US-10-424-599-32675 (1-2586)

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 28, 2005, 02:20:58 ; Search time 109 Seconds

(without alignments)
2373.644 Million cell updates/sec

Title: US-09-886-400A-4

Perfect score: 1879

Sequence: 1 LRALVPHGNLQYAEIPKSEI.....RLDPAFRALYNDMGENCEP 364

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n_model -DEV=xlp
-Q=/cgn2_1/USPRO_epool_p/US09886400/runat_27012005_055722_17947/app_query.fasta_1.519
-DB=Issued_Patents_NA -QEMT=fastcap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886400 @cgn2_1_105 @runat_27012005_055722_17947 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*\n2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*\n3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*\n4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*\n5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*\n6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|-------|---------------------|
| 1 | 1879 | 100.0 | 1095 | 4 | US-10-166-606-3 |
| 2 | 1879 | 100.0 | 1095 | 4 | US-10-112-231A-3 |
| 3 | 1589 | 84.6 | 1041 | 2 | US-08-613-220B-3 |
| 4 | 146.5 | 7.8 | 1664976 | 4 | US-08-916-421B-1 |
| 5 | 146.5 | 7.8 | 1664976 | 4 | US-09-692-570-1 |
| 6 | 140.5 | 7.5 | 3139 | 1 | US-07-894-212A-1 |
| 7 | 105 | 5.6 | 684 | 4 | US-09-107-532A-1900 |
| 8 | 104 | 5.5 | 4403765 | 3 | US-09-103-840A-2 |
| 9 | 104 | 5.5 | 4411529 | 3 | US-09-103-840A-1 |
| 10 | 103 | 5.5 | 5030 | 4 | US-08-956-171E-324 |
| 11 | 103 | 5.5 | 5030 | 4 | US-08-781-986A-324 |
| 12 | 102.5 | 5.5 | 1912 | 1 | US-08-270-013B-1 |

| | | | | | | |
|----|-------|-----|---------|---|----------------------|----------------------|
| 13 | 102.5 | 5.5 | 1912 | 1 | US-08-838-418-1 | Sequence 1, Appli |
| 14 | 101.5 | 5.4 | 1083 | 4 | US-09-710-279-2149 | Sequence 2149, Ap |
| 15 | 101.5 | 5.4 | 1131 | 4 | US-09-710-279-2105 | Sequence 2105, Ap |
| 16 | 101.5 | 5.4 | 3014 | 4 | US-09-710-279-4330 | Sequence 4330, Ap |
| 17 | 101.5 | 5.4 | 3041 | 4 | US-09-710-279-4336 | Sequence 4368, Ap |
| 18 | 99 | 5.3 | 536165 | 4 | US-09-214-808-1 | Sequence 1, Appli |
| 19 | 98.5 | 5.2 | 1800 | 4 | US-09-614-221A-209 | Sequence 209, App |
| 20 | 97 | 5.2 | 1962 | 4 | US-09-540-286-1208 | Sequence 1208, Ap |
| 21 | 97 | 5.2 | 100848 | 4 | US-09-596-002-39 | Sequence 39, Appli |
| 22 | 94.5 | 5.0 | 1104 | 3 | US-09-114-001C-841 | Sequence 841, App |
| 23 | 94.5 | 5.0 | 3183 | 4 | US-09-221-017B-939 | Sequence 939, App |
| 24 | 93.5 | 5.0 | 100990 | 4 | US-09-409-800B-2 | Sequence 2, Appli |
| 25 | 92 | 4.9 | 1128 | 4 | US-09-710-279-1839 | Sequence 1839, Ap |
| 26 | 92 | 4.9 | 1449 | 4 | US-09-252-991A-15134 | Sequence 15134, A |
| 27 | 92 | 4.9 | 2226 | 4 | US-09-252-991A-14877 | Sequence 14877, A |
| 28 | 92 | 4.9 | 2268 | 4 | US-09-252-991A-14487 | Sequence 14487, A |
| 29 | 92 | 4.9 | 2538 | 4 | US-09-248-796A-1291 | Sequence 1291, Ap |
| 30 | 92 | 4.9 | 4086 | 4 | US-09-710-279-1765 | Sequence 3765, Ap |
| 31 | 92 | 4.9 | 1664976 | 4 | US-08-916-421B-1 | Sequence 1, Appli |
| 32 | 92 | 4.9 | 1664976 | 4 | US-09-692-570-1 | Sequence 1, Appli |
| 33 | 91.5 | 4.9 | 3030 | 4 | US-09-693-146-3 | Sequence 3, Appli |
| 34 | 91.5 | 4.9 | 3113 | 4 | US-09-894-998A-52 | Sequence 52, Appli |
| 35 | 91.5 | 4.9 | 3345 | 4 | US-09-894-998A-49 | Sequence 49, Appli |
| 36 | 91.5 | 4.9 | 3350 | 4 | US-09-894-998A-48 | Sequence 48, Appli |
| 37 | 91.5 | 4.9 | 5421 | 1 | US-08-118-441-28 | Sequence 28, Appli |
| 38 | 91.5 | 4.9 | 5421 | 3 | US-08-338-579A-28 | Sequence 28, Appli |
| 39 | 91.5 | 4.9 | 5421 | 5 | PCT-US94-09851-28 | Sequence 28, Appli |
| 40 | 91 | 4.8 | 3096 | 4 | US-09-543-681A-4073 | Sequence 4073, Appli |
| 41 | 90.5 | 4.8 | 1221 | 3 | US-09-134-001C-733 | Sequence 733, App |
| 42 | 90.5 | 4.8 | 2370 | 4 | US-09-248-796A-5191 | Sequence 5191, Ap |
| 43 | 90.5 | 4.8 | 3516 | 4 | US-09-248-796A-1940 | Sequence 1940, Ap |
| 44 | 90 | 4.8 | 4074 | 4 | US-09-252-991A-4737 | Sequence 4737, Ap |
| 45 | 90 | 4.8 | 4092 | 4 | US-09-252-991A-4771 | Sequence 4771, Ap |

ALIGNMENTS

RESULT 1
US-10-166-606-3
; Sequence 3, Application US/10166606
; Patent No. 6644756
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/166, 606
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 09/407, 806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 08/613, 220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alcaliphilus
US-10-166-606-3

Alignment Scores:
Pred. No.: 1,16e-230 Length: 1095
Score: 1879.00 Matches: 364
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-886-400A-4 (1-364) x US-10-166-606-3 (1-1095)
Oy 1 LeuArgLauValPheHisGlyAsnLeuGlnTyrIaGluIleProLysSerGluIle 20
|||||

Db 1 TTGAGAGCGCTGCTTTTCACGGCAACCTCCAGTATGCCAAATCCCAAGAGCGAAATC 60
Qy 21 ProlysvallilegualyvalaThyrilleProvalilegultThrleuilelysgluile 40
Db 61 CCANAGCTCATAGAGAGCATACATCCAGTCAATCGAGACACGATTTAAAGGAATTT 120
Qy 41 ProPheglileuAanileThrglyThyrThreulysPheleuProlysaPileasp 60
Db 121 CCTTTGGGCTCAACATAGAGGGCATATACCTTTAAAGTTCCCTCCGAAGATTTATAGAC 180
Qy 61 leuVallysglygyllealeSeraSpleuilegluileilegylThrSerTyThHis 80
Db 181 CTCCTTAAGGGGGCATCGCAGTGCATGATAGATTAATCGGAACGAGCTACACCCAC 240
Qy 81 AlalleuProleuProleuSerArgvalaglualeglnvalaglnaArgAspArglu 100
Db 241 GCATATCTCCCTCTGCTCCGCTTACAGAGATAGACACAAAGTTACAGAGATAGGAA 300
Qy 101 VallysgluileuPhegluValSerProlysglyPheTrpleuProgluLeuAlaTy 120
Db 301 GTTAAGAGAGAGCTCTTCAGAGTTCTCCAAAGGATTCGGCTGCCAGAGCTGCCTAT 360
Qy 121 AspProleilePProalalleleuLyasAsnglyTyrglyTyreulePhealaspGly 140
Db 361 GACCCGATATCCCTGCTCATACGAAAGACAAACGGTATAGATATCTATTCGCCAGCGG 420
Qy 141 GlualameteuPheSeraAlaileuAsnSeraAlaileysProilleysProleuTy 160
Db 421 GAGCGATGCTTTCTCAGCTCATCTCACTCGCGCATTAAGCCAAATTAAGCGCTTAT 480
Qy 161 ProHisleuileysalaglnaArggluLyasArgPheArgTyrlleSeryLeuLeugly 180
Db 481 CCACACCTTAAAGGCCCAAGGAAAGCGCTTATGATACATCAGCTATCTCTTGCT 540
Qy 181 leuArggluileuArglysalallelyleuValPhegluLyysValThrleuLyala 200
Db 541 CTCAGGAGCTTGAAGAGCGATTAAGCTGCTTTTGAAGATTAAGTAAAGCTTAAAGCGCA 600
Qy 201 VallysaPileglualaValProvalTrpValAlaValaAnthraAlaValameteuLeugly 220
Db 601 GTCAAGACATCGAAGCCGATACCGCTTGGGTGGCGGTGAACACGGCTGTAAGCTCGGC 660
Qy 221 lilegylArgleuProleuMetanProlysaValaIasertPilegluasp 240
Db 661 ATCGAAGGCTTCTCTTATGAATCTTAAGAAAGTGGCGAGCTGATAGAGCAAGAC 720
Qy 241 AsnilleuLeuTyrglyThyrAspillegulePheilegylTyraArgaspilleaglyTy 260
Db 721 AACATTTCTTATACGGCACCGATTAAGATTCATTGGCTATAGGACATTTGACAGGCTAC 780
Qy 261 ArgMetSerValgluileuLeuGluValileaspGluLeuAsnSerGluLeuCySleu 280
Db 781 AGATGAGGTGAGGGATTTATAGAGTTATAGAGCTCAACTCGAATCTGCTCTT 840
Qy 281 ProSerGluLeuLyshisSerglyArggluLeuTyreulysThrSerTrpAlapro 300
Db 841 CCTTCAGAGCTGAAGCACGTGAGAGGAGCTCTACTTACGACCTTCAGATTGGCACCA 900
Qy 301 AspLySerleuArglyleTPaArgluAspGluGluAsnAlaArgleuAsnmetLeuSer 320
Db 901 GAATAGAGCTTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 321 TyraMetSerArglygluLeuAlaPheleuAlaGluAsnSerAspAlaArglyTyrglu 340
Db 961 TACAAATATAGAGGGCGAATCGCTTTTATAGCCGACAAACGATCCAAAGGGATAGGAG 1020
Qy 341 ProleuProgluArgArgleuAspAlaPheArgAlaileTyraAsnAspTrpArgglu 360
Db 1021 CCCCTCCCTGAAGAGAGGCTGATGCTTCCGGCGATATATTAAGATTTGAGGGGTGA 1080
Qy 361 AsnglygluPro 364
Db 1081 AATGGGAACT 1092

RESULT 2
US-10-112-231A-3
; Sequence 3, Application US/10112231A
; Patent No. 6744246
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE OF INVENTION: MAKING AND USING THEM (Amended)
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/112,231A
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: US 09/407,806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Thermococcus alkaliphilus
US-10-112-231A-3

Alignment Scores:
Pred. No.: 1,16e-230 Length: 1095
Score: 1879.00 Matches: 364
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-886-400A-4 (1-364) x US-10-112-231A-3 (1-1095)
Qy 1 leuArgAlaLeuValPheHilegylAsnleuGlnTyraGluileProlysaSergluile 20
Db 1 TTGAGAGCGCTGCTTTTCACGGCAACCTCCAGTATGCCAAATCCCAAGAGCGAAATC 60
Qy 21 ProlysvallilegualyvalaThyrilleProvalilegultThrleuilelysgluile 40
Db 61 CCANAGCTCATAGAGAGCATACATCCAGTCAATCGAGACACGATTTAAAGGAATTT 120
Qy 41 ProPheglileuAanileThrglyThyrThreulysPheleuProlysaPileasp 60
Db 121 CCTTTGGGCTCAACATAGAGGGCATATACCTTTAAAGTTCCCTCCGAAGATTTATAGAC 180
Qy 61 leuVallysglygyllealeSeraSpleuilegluileilegylThrSerTyThHis 80
Db 181 CTCCTTAAGGGGGCATCGCAGTGCATGATAGATTAATCGGAACGAGCTACACCCAC 240
Qy 81 AlalleuProleuProleuSerArgvalaglualeglnvalaglnaArgAspArglu 100
Db 241 GCATATCTCCCTCTGCTCCGCTTACAGAGATAGACACAAAGTTACAGAGATAGGAA 300
Qy 101 VallysgluileuPhegluValSerProlysglyPheTrpleuProgluLeuAlaTy 120
Db 301 GTTAAGAGAGAGCTCTTCAGAGTTCTCCAAAGGATTCGGCTGCCAGAGCTGCCTAT 360
Qy 121 AspProleilePProalalleleuLyasAsnglyTyrglyTyreulePhealaspGly 140
Db 361 GACCCGATATCCCTGCTCATACGAAAGACAAACGGTATAGATATCTATTCGCCAGCGG 420
Qy 141 GlualameteuPheSeraAlaileuAsnSeraAlaileysProilleysProleuTy 160
Db 421 GAGCGATGCTTTCTCAGCTCATCTCACTCGCGCATTAAGCCAAATTAAGCGCTTAT 480
Qy 161 ProHisleuileysalaglnaArggluLyasArgPheArgTyrlleSeryLeuLeugly 180
Db 481 CCACACCTTAAAGGCCCAAGGAAAGCGCTTATGATACATCAGCTATCTCTTGCT 540
Qy 181 leuArggluileuArglysalallelyleuValPhegluLyysValThrleuLyala 200

Db 541 CTCAGGAGCTTAGAGAGCGATTAAGCTCTGTTTGAAGTAAAGTAAAGCTTAAGCA 600
 Oy 201 Vallysaapilleglualavalprovaltrpvalalavalantthralavalmetleugly 220
 Db 601 GTCAAGACATCGAAGCCGATCCCTTTGGGGCGGAGAACACGGCTGTATATGCTCGGC 660
 Oy 221 lleglyargleuproleumetasepvalalasetrplleglyuaplyasap 240
 Db 661 ATCGAAGGCTTCCCTTATATGATCTTAAGAAAGTGGGAGCTGATAGAGACAAAGAC 720
 Oy 241 Asnilleuleuetyrglythraspilleglubheilleglytrargaspilleagllytr 260
 Db 721 AACATTTCTTATACGGCACCGCATATAGATTCAATGGCTATAGGACATTCAGGCTAC 780
 Oy 261 ArgsetserValgluglyleuleuglyuvalilleaspgluleuasepsergluleucyaleu 280
 Db 781 AGAATGAGGTGTGAGGATATATAGAGGTATATAGCGAGCTCAACTCGAACTGTGCTTC 840
 Oy 281 Prosergluleuylshiserglyarggluleuetyrleuargthsersestrpallapro 300
 Db 841 CCTCAGAGCTGAAGACACAGTGAAGGAGCTCTTACGACCTTCAGATTGGGACACA 900
 Oy 301 Asplyserleuargyletrparggluapgluglyaenallargyleuasepser 320
 Db 901 GATTAAGAGCTTAGAGATATGAGAGAGACGAAAGGAGACGCAAGCTTATATATGCTGTCC 960
 Oy 321 Tyrasmetaragllygluleuualaphelenuaglunaseraspallarglytrpglu 340
 Db 961 TACATATGAGGGGAGAACTGCTTTTACCCGAAACAGCGAAGCAAGGGATGGGAG 1020
 Oy 341 Proleuproglyuargargleuaspallapheargalaileryrasnasprrparglyglu 360
 Db 1021 CCCCTCCCTGAGAGAGGCTGATCCCTTCGGGCGATATATACGATTGAGGGGTGAA 1080
 Oy 361 Asnlyglupro 364
 Db 1081 AATGGGAGACCT 1092

RESULT 3

US-08-613-220B-3
 ; Sequence 3, Application US/08613220B
 ; Patent No. 5958751
 ; GENERAL INFORMATION:
 ; APPLICANT: Murphy, Dennis
 ; APPLICANT: Reid, John
 ; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/613,220B
 ; FILING DATE: 08-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hallie, Ph.D., Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 09010/004001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-678-5070
 ; TELEFAX: 619-68-5099

; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1041 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 1...1038
 ; OTHER INFORMATION:
 ; US-08-613-220B-3
 Alignment Scores:
 Pred. No.: 1.55e-193 Length: 1041
 Score: 1589.00 Matches: 346
 Percent Similarity: 95.05% Conservative: 0
 Best Local Similarity: 95.05% Mismatches: 0
 Query Match: 84.57% Indels: 18
 DB: 2 Gaps: 18
 US-09-886-400A-4 (1-364) x US-08-613-220B-3 (1-1041)
 Oy 1 LeuarglaleuValPheHieglYAsnleugIntYrAlagluileProlySergluile 20
 Db 1 TTGAAGCGCTCGCTCTTTCACGCGACACTCCAGTATGCCAAATCCCAAGAGCGAA 57
 Oy 21 ProlyValillegluYsalaYrileProvalillegluThrleuileYsgluile 40
 Db 58 CCAAGGTCAATAGAGAGGCAATACATCCAGTCAGACATCGATTAAAGAGAA 114
 Oy 41 ProheglYleuanillethnglyrThrleuylsPheleupProlySaspilleleasp 60
 Db 115 CTTTGGGCTCAACATACAGGGCTATACCTTAAAGTTCTCCCGAAGGATATTATA 171
 Oy 61 leuValylsglyglYileAlaseraspLeuilegluileleglYThrserYrThHis 80
 Db 172 CTCGTTAAAGGGGCAATCGGAGTGAACCTGATAGATATATGGAACGAGCTACACG 228
 Oy 81 AlalileupProleuProleuSeraYrValgluAlagluValAlnarGasparGlu 100
 Db 229 GCATATCTCCCGCTCCCTGCGGCTTAGAGAGTGAACACAAAGTTAGAGATAGG 285
 Oy 101 VallysegluYleuphegluValSerProlysglyPheTrleupProgluLeuAlaYr 120
 Db 286 GTTAAGAGAGGCTCTTCGAGTTTCCTCAAGGGATTCTGCGCCAGAGCTGCC 342
 Oy 121 AspprolleleproAlalileuYsaspAsnlyrYrgluYrleupheAlaAspGly 140
 Db 343 GACCGATTAATCCCTGCCATATCTGAAGACAAACGGTTATAGATATATTTCGCCGAC 399
 Oy 141 GluAlametrleupheseralAlisLeuasepAlalleyspProilleyspProleuYr 160
 Db 400 GAGCGATGCTTTTCTCAGCTCATCTCACTCGCGATTAAGCAATTAAACCGCTC 456
 Oy 161 ProHieuleuileYsAlaglnarggluYsarpheargTrilleSeryrleuLeugly 180
 Db 457 CCACACCTTATTAAGGCCCAAGGAAACCGCTTAGTACATCACTACTCTCT 513
 Oy 181 leuarggluYleuargYsAlalleYseuValPhegluYlyysValThrleuYsAla 200
 Db 514 CTCAGGAGCTTAGAAGGCGATTAACCTGTTTGAAGGTAAAGTAAAGCTTAAG 570
 Oy 201 Vallysaapilleglualavalprovaltrpvalalavalantthralavalmetleugly 220
 Db 571 GTCAAGACATCGAAGCCGATCCCTTTGGGTGGCCGTGAACACGGCTGTAAATGCTC 627
 Oy 221 lleglyargleuproleumetasepvalalasetrplleglyuaplyasap 240
 Db 628 ATCGAAGGCTTCCCTTATGAAATCCTAAGAAAGTGGCGAGCTGATAGAGACAAG 684
 Oy 241 Asnilleuleuetyrglythraspilleglubheilleglytrargaspilleagllytr 260

Db 685 AACATCTCTATACGGCACCGATATAGATTGATGGCTATAGGACATTGACGC--- 741
Qy 261 ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCyLeu 280
Db 742 AGAATGATGTGTGGAGGATTTATTAAGAGTTATACACAGCTCACTGGAACTGTG--- 798
Qy 281 ProSerGluLeuLysHisSerGlyArgGluLeuTyrLeuArgTyrHisSerTTPalaPro 300
Db 799 CCCTCAGAGCTGAGGACACAGTGSAGGAGGAGCTCACTTACGAGCTTGAAGTTGGGC--- 855
Qy 301 AspLysSerLeuAlaGlyIleTTPArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSer 320
Db 856 GATTAAGGCTTGAAGATATGAGAGAGACGAGGAGAACCGCAAGACTTAATATGCTG--- 912
Qy 321 TyrAsnMetArgGlyGlyLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTTPGlu 340
Db 913 TACATATGAGGGGCGAAGCTGCTTTTAAAGCCAGAAACGAGTGCAGAGGGATGG--- 969
Qy 341 ProLeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTTPArgGlyGlu 360
Db 970 CCCCTCCCTGAGAGGAGGCTGAGATGCTTCCGGGCGATATATACGATTTGAGGGGT--- 1026
Qy 361 AsnGlyGluPro 364
Db 1027 AATGGGGAACCT 1038

RESULT 4

US-08-916-421B-1/C
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Buit et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ. ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ. ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g

```

NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1115881)..(1115881)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, c, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, c, c, or g
US-08-916-421B-1

```

```

Alignment Scores:
Pred. No.: 0.00235 Length: 1664976
Score: 146.50 Matches: 87
Percent Similarity: 40.58% Conservative: 66
Best Local Similarity: 23.08% Mismatches: 129
Query Match: 7.80% Indels: 95
DB: 4 Gaps: 24

```

```

US-09-886-400a-4 (1-364) x US-08-916-421B-1 (1-1664976)
QY 14 GluileProlySerGluileProlyValileGluValiIleProValileGlu 33
DB 1550843 AAATTAATAAGAGATTGTTTAAATTAAGTGGCTAATAATGCTACATTCCTAACAAATGAG 1550784

```

```

QY 34 Thrileuilelys-----GluGluileProPheGlyLeuAsn-----IleThrGly--- 48
DB 1550783 TTGATATTGACGCTTATTGAGTAATATTTAAAGTAACTTATTCATTTCTGGGCT 1550724
QY 49 -----TyrThrleuLysPheLeuProLysAsp-----IleIleAspLeuValileGly 64
DB 1550723 TTTGAGAGCAAGCTTTGGAATTT-----AATGATTCGGTGGATTATTTAAAGAT 1550670
QY 65 GlyIleAlaSerAspLeuileGluileIleleGlyThrSerTyrThrIleAlaIleLeuPro 84
DB 1550669 TTGGTAAAGACTGGCAATGTGAGTTGATGCTGAAACATATTCATCTCACTAACAACT 1550610
QY 85 LeuileuProleu-----SerArgValGluValleGluValGlnValGlnValGlnValGlu 103
DB 1550609 CTATTGTAAGTGAAGATGAGTTTATTTGAAGATTTGATGATGATGATGATGATGATGATGAT 1550550
QY 104 GluileuPheGluValSerProLysGlyPheTyrleuProGluValleAlaTyrAspProIle 123
DB 1550549 GAAATATTGTTGTTTAAAGCAAGGTTTGAATACTGAGCTATATACAAACAGAA 1550490
QY 124 IleProAlaIleleuLysAspAsnGlyTyrGluTyrleuPheAlaAspGly---GluAla 142
DB 1550489 ATTGCAAGATAGCAAAAGATTTAGGGTTTAAAGCATATTACTGAGGGAATTGAGAAA 1550430
QY 143 MetleuPheSerAlaIleLeuAsnSerAlaIleLysProIleLysProLeuTyrProHis 162
DB 1550429 ATCTTA----- 1550424
QY 163 LeuileuLysAlaGlnArgGluLysArgPheArgTyrIleSer-----TyrleuLeu 179
DB 1550423 -----GGCTGAGGCTCCCAACATCTTTTACCATCACCAGATGATGAAATTT 1550373
QY 180 GlyleuArgGluLeuArgLysAlaIleLysLeuValPheGluGlyLysValThrleuLys 199
DB 1550372 CTGTTAAGCAATTTAGGTTGAGTACGATGATGCTGCTTTAGA-----TTT 1550328
QY 200 AlaValLysAspIleGluAlaValPro-----ValTyrValAla 212
DB 1550327 TCAGCAAGATTTGGGACCAATTCCTTAACAGCTGATTAATTCCTATTGTTGTTACT 1550268
QY 213 ValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLysLeuVal 232
DB 1550267 TCACTCTCGGGGAAGTTATTAACATA-----TATATGACATAGACACATTT 1550220
QY 233 Alaser-----TyrIleGluAspLysAspAsnIleLeuLeuTyrGlyThrAspIle--- 249
DB 1550219 GGAGAACATCACTGAAAGAA-----ACTGAAATATTT 1550187
QY 250 GluPheIleGlyTyrArg-----AspIleAlaGlyTyr---ArgMetSerValGluGly 266
DB 1550186 GAGTTTGTGAGATTTTGGCTTATGAGATACATGACATGAACTTTAGAGGTTGTTAAT 1550127
QY 267 LeuLeuGluValIleAspGluLeuAsnSerGluLeuProSerGluLeuLysHis 286
DB 1550126 GTCACTGAGGTTGTGATAGTTTGAAGCAAGA----- 1550094
QY 287 SerGlyArgGluLeuTyrLeuArg-----ThrSerTyrAla---ProAspLys 302
DB 1550093 -----GGAGAGATTTACGTGCAATGATTTGCCACATATCATCGGCTGACACAGAGAG 1550040
QY 303 SerleuArgIleTyrPheArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSerTyrAsn 322
DB 1550039 GATGTAAAGTGCATGTTA-----GGGAAAT---AAGATGCAAGAAATTCGTTGAA 1549992
QY 323 MetArgGlyLeuLeuAla---PheLeuAlaGluLysSerAspAlaArgGlyTyrGluPro 341
DB 1549991 AAACATAAGATATGTTAAATTTATTAAGAGAACTCAAACTTA----- 1549944
QY 342 LeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsnAspTyrArg 358
DB 1549943 -----AAAAAATTAAATTTGATGCAAAATATATAGATGTATTAAG 1549902

```

RESULT 5

US-09-692-570-1/c
Sequence 1, Application US/09692570
Patent No. 6797466
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6797466
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275c1
CURRENT APPLICATION NUMBER: US/09/692,570
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
PRIOR APPLICATION NUMBER: US 08/916,421
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g

FEATURE:
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:


```
/ GENERAL INFORMATION:
/ APPLICANT: ASADA, KIYOZO
/ APPLICANT: UEMORI, TARASHI
/ APPLICANT: MUKAI, HIROYUKI
/ APPLICANT: KATO, IKUNOSHIN
/ APPLICANT: LADERMAN, KENNETH
/ APPLICANT: ANFINSEN, CHRISTIAN
/ TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CUSHMAN DABRY & CUSHMAN
/ STREET: 1100 NEW YORK AVENUE, N.W.
/ CITY: WASHINGTON, D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/894,212A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KOKULIS, PAUL N.
/ REGISTRATION NUMBER: 16773
/ REFERENCE/DOCKET NUMBER: 95469/C-1195
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 861-3000
/ TELEFAX: (202) 822-0944
/ TELETYPE: 6714627 CUSH
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3139 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-07-894-212A-1

Alignment Scores:
Pred. No.: 4,09e-07 Length: 3139
Score: 140.50 Matches: 89
Percent Similarity: 38.11% Conservative: 60
Best Local Similarity: 22.76% Mismatches: 127
Query Match: 7.48% Indels: 115
DB: 1 Gaps: 23

US-09-886-400A-4 (1-364) x US-07-894-212A-1 (1-3139)

QY 25 GtUuYAlAtYrIlleProValIleGluThrLeuIleLysGluGluPro----- 41
DB 800 GAAAGGTATGCGACCGCTTTCGAGACTCTG-----GAGGAATATCCAAACATGAAG 853
QY 42 PheGlyLeuAsnIleThrGlyYrThrLeuYpHeLeuPro-----LysAspIle 58
DB 854 GTTGCACTCATACAGTGGCCCCCTCATGAGTGCCTCAATAATAGACCCGAATAC 913
QY 59 IleAspLeuValIleGlyYrIleAlaSerAspLeuIleGluIleIleGlyThrSerYr 78
DB 914 ATAGACTTCTTGAAGTCTAGGAAAGAGACAGAGTGAAGTACGTCGTCGGGCTTC 973
QY 79 ThrHAlaIleLeuProLeuSerArgValGluAlaGlnValGlnAgaGAP 98
DB 974 TAGACACCTGTGTACATCAATCCCA-----AAGGAGATGAAATAGACAGATA 1024
QY 99 ArgGluValIleGlu-----GluLeuPheGluValSerProLysGlyPheTrpLeuPro 116
DB 1025 AGGTATATGAAAGAGTGGCTAAGATGATTTGATGCTAGGAGAGTTGGCTAACT 1084
QY 117 GluLeuAlaIleYrAspProIleIleProAlaIleLeuLysAspAsnGlyYrGluYrLeu 136
DB 117 GtUuYAlAtYrAspProIleIleProAlaIleLeuLysAspAsnGlyYrGluYrLeu 136
```

```
DB 1085 GAAAGGTATGCGACCGCTTTCGAGACTCTTAAAGACCGGAATAGATTATGTA 1144
QY 137 PheAlaAspGlyGluAlaMetLeuPheSerAlaIleLeuAsnSerAlaIleLysProIle 156
DB 1145 ATAGTTGAC-----GATTACCACTTCAATGAGTGGGATTAAGTAA 1186
QY 157 LysProLeuYr---ProHsIleuIleLysAlaGlnArgIle----- 169
DB 1187 GAGAGCTGTACGCGCATATATACGAAGATGTGGGAGATTATAGCTGTTTCCCG 1246
QY 170 -----LysArgPheArgYrIleSerYrLeuLeuGlyLeuArgIleuArgLysAla 187
DB 1247 ATAGATGAGAAAGTTGAG-----TATTGATTCCTTTAGACCCGTGATAGGTC 1297
QY 188 IleLysLeu-----ValPheGluGlyLysValThrLeuLysAlaVal-----Lys 202
DB 1298 TTGAATATACCTGCACTCTCATAGATGTGATGAGACGAAGTTGCAGATTTTCATGAC 1357
QY 203 AspIleGluAlaValProValIleProValAlaValAsnThrAlaValMetLeuGlyIleGly 222
DB 1358 GATGTGAGAAAGTTTGAATCTGGCCTGGAACCTTATGATGAGTGGGTAT----- 1405
QY 223 ArgLeuProLeuMetAsnProLysValAlaSerTrpIle----- 236
DB 1406 -----GAAAAGGATGTTTAAAGAAATTTCTTGATAGA 1438
QY 237 -----GluAspLysAspAsnIleLeuLeuYrGlyYrThrAspIleGlu----- 250
DB 1439 ATTTCAAGTATGAAAGATAAAGTAACTTAATGCTTTACAGTAATCTTACAAAATATAG 1498
QY 251 -----PheIleGlyYrThrAspIleAlaGlyYrArgMetSerValGluLysLeuLeu 268
DB 1499 CCTAGAGCTTGTGTTATTTTCCAAATGCTTCAAT-----TTT 1537
QY 269 GluValIleAspGluLeuAsnSerGluLeuCysLeuPro----- 281
DB 1538 GAGATG-----AGCAATGCTATTGTCAGCAAGACGACGACCTTT 1582
QY 282 -----SerGluLeuLysHisSerGly-----ArgGluLeuYrLeu 293
DB 1583 GTGAGTTGTCAAAGAGCTTAAAGTTAAAGTATATTTGAAAAGTACAGGATATTTGTT 1642
QY 294 ArgThrSerSerTrpAlaProAspLysSerLeuArgIleTrpArgGluAspIleGlnLys 313
DB 1643 AGCGAGGAATTTG-----AAGATTTTCTCTATTAATATCCCAAGACGACAC 1690
QY 314 -----AlaArgLeuAsnMetLeuSerYrAsnMetArgGlyGluLeuAlaPheLeu 330
DB 1691 TACATGACAAAGAAATGCTATGTTAGTTAAGTTAGAGAAACAA----- 1738
QY 331 AlaGluAsnSerAspAlaArgGlyYrProIleProLeuProGluArgArgLeuAspAlaPhe 350
DB 1739 -----CCTGAGGCCAGGAAGATATCTGCTG 1762
QY 351 ArgAlaIleYrAsnAsp-----TrpArgGly 359
DB 1763 AGAGCAATATGATGATGCTTATTGGACGCGC 1795

RESULT 7
US-09-107-532A-1900
Sequence 1900, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
```



```

1      ZIP: 02354
2
3      COMPUTER READABLE FORM:
4
5      MEDIUM TYPE: CD/ROM ISO9660
6
7      COMPUTER: PC
8
9      OPERATING SYSTEM: <Unknown>
10     SOFTWARE: ASCII
11
12     CURRENT APPLICATION DATA:
13
14     APPLICATION NUMBER: US/09/107,532A
15
16     FILING DATE: 30-Jun-1998
17
18     PRIOR APPLICATION DATA:
19
20     APPLICATION NUMBER: 60/085,598
21
22     FILING DATE: 14 MAY 1998
23
24     APPLICATION NUMBER: 60/051571
25
26     FILING DATE: July 2, 1997
27
28     ATTORNEY/AGENT INFORMATION:
29
30     NAME: Ariadello, Pamela Deneke
31
32     REGISTRATION NUMBER: 40,489
33
34     REFERENCE/DOCKET NUMBER: GTC-012
35
36     TELECOMMUNICATION INFORMATION:
37
38     TELEPHONE: (781)893-5007
39
40     TELEFAX: (781)893-8277
41
42     INFORMATION FOR SEQ ID NO: 1900:
43
44     SEQUENCE CHARACTERISTICS:
45
46     LENGTH: 684 base pairs
47
48     TYPE: nucleic acid
49
50     STRANDEDNESS: double
51
52     TOPOLOGY: circular
53
54     MOLECULE TYPE: DNA (genomic)
55
56     HYPOTHETICAL: NO
57
58     ANTI-SENSE: NO
59
60     ORIGINAL SOURCE:
61
62     ORGANISM: Enterococcus faecium
63
64     FEATURE:
65
66     NAME/KEY: misc feature
67
68     LOCATION: (b) LOCATION 1...684
69
70     SEQUENCE DESCRIPTION: SEQ ID NO: 1900:
71
72     US-09-107-532A-1900

```

| | |
|------------------------|--------|
| Alignment Scores: | |
| Pred. No.: | 0.0015 |
| Score: | 105.00 |
| Percent Similarity: | 37.90% |
| Best Local Similarity: | 24.19% |
| Query Match: | 5.59% |
| DB: | 4 |
| Gaps: | 12 |
| Length: | 664 |
| Matches: | 105 |
| Conservative: | 34 |
| Mismatches: | 88 |
| Indels: | 66 |
| Gaps: | 12 |

US-09-886-400A-4 (1-364) X US-09-107-532A-1900 (1-684)

[illegible]

```

Db      292  TCTGTTCTATCAAAAGAC-----GGCGCATCAACATGACAGAGTTACCAAAATAC 342
      +
      ::::|||||:::
Oy      261  ArgMetSerValGluGlyLeuLeuGluValIleAspGluLeuAsnSergLyuCyLeu 280
      ::::|||||:::
Db      343  CATATCGAGATGAGACGATGGCGACACATC-----TTAAGACAGCAAGAAATGCTTT 397
      ::::|||||:::
Oy      281  ProSergLyu-----LeuIlyHisSergIlyArgGluLeuTyrLeuArg 294
      |||
      ::::|||||:::
Db      394  TTTCTTGAAATATCCAAACATGTTATCTTACAGAAATGCCGATCAAAATGATGTAATGAAA 453
      ::::|||||:::
Oy      295  ThrSergTrrPalProAspLysSergLeuArgIleTrrArgGluAspGluGlyAsnAla 314
      |||
      ::::|||||:::
Db      454  AAATCAGAT-----GAAAGACAAACCTTATATGATCCGTGTTAGATGAAAGGCAAAATT 507
      ::::|||||:::
Oy      315  ArgLeuAsnMetLeuSergTrrAsnMetArgIlyGluLeuAlaPheLeuAlaGluAsnSer 333
      ::::|||||:::
Db      508  CACACATAAAGTGGCTTCATCGTCCAAATCAATTGACTGAA---GCATGGCTTATGAGAACCTTG 564
      ::::|||||:::
Oy      335  AspAlaArgGlyTrrGluLyrProLeuProGluArgAlaGluAspAlaPheAsnArgAlaIleTyr 354
      |||
      ::::|||||:::
Db      565  AAGAAAGAAAGGCTAT-----GCAGATGTAAACACACGCTGATTTAT 603
      ::::|||||:::
Oy      355  AsnAspTrrArgGlyGluAsnGly 362
      ::::|||||:::
Db      604  GTTAAATGGTCAGAAAGAAAAGGG 627

```

RESULT 8

; Sequence 2, Application US/09103840A

GENERAL INFORMATION:

APPLICANT: WHITE, Owen R.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: TUBERCULOSIS ;

CURRENT APPLICATION NUMBER: US/09/103,840A

NUMBER OF SEO ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1
; SEO ID NO 2

```

; LENGTH: 4403765
; TYPE: DNA

```

ORGANISM: *Mycobacterium tuberculosis*
FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Alignment Scores:

Score:

Best Local Similarity

query match:
DB:

US-09-886-400A-4 (1-364) X US-09-103-840A-2 (1-4403765)

61 LeuValIwG]vgIvIleAlSeRAspLeuIlegIvIleIlegIvThrSerTyvThrHis 80

3386055 CTGCTGCGCGCTGATTCAGCGCGGACAGGTGGAGCTGCTCGGTGCGCCACCTTTCGCGCAC 3386114

81 A]at]el,eu]Pro]e]u]SerA]rva]G]u]A]aG]nva]G]nA]rca]a]nA]rG]u 100

[illegible]

101 *W3T::C3H/10T1/2ClPhoc3::W3T::C3H/10T1/2ClPhoc3* 114

[illegible]

```

Db      3386157 CTGCGGAGAAAGCCCTCGCGATGCTAGCTCGGCTGGCGCACCCGAAAGGATCTGG 3386216
;
Qy      115 LeuProGluLeuAlaTyrAspProIleIleProAlaIleLeuLysAspAngLYTyrGlu 134
;
Db      3386217 GCACCCGAATGCCGATACGCCGCCGGAGATGAGTGCATACGCCACCGGGAGTCACT 3386276
;
Qy      135 TyrLeuPheAlaAspGlyGluAlaMetLeuPheSerLahIleuAsnSerAlaIle--- 153
;
Db      3386277 CACTTCATGCTGCACGGCCCG-----TCGCTGACCGCGCACACCGCGCTGGCG 3386324
;
Qy      154 LysProIle----- 156
;
Db      3386325 CGGCGGTGGGGAACCGATGCTGCGCTTCGCTGCGACCTTGAGAGTCACTACCGG 3386384
;
Qy      157 -----LysProLeuTyrPro---HisLeuIleLysAlaGlnArgGluLysArg 171
;
Db      3386385 GTGTGTCACCCGAATCCGGCTACCCCGGGCAC-----GCCCTTACCGCGAC----- 3386432
;
Qy      172 PheArgTyrIleSerTyrLeuLeuGlyLeuArgGluLeuArgLysAlaIleLysLeuVal 191
;
Db      3386433 TTCACACCTACGACACCTGACCGGATCAACCGGCGGCGTCAACCGGCGCTAACGTC 3386492
;
Qy      192 PheGluGlyLysValThrLeuLysAlaValLysAspIleGluAlaValProValTyrVal 211
;
Db      3386493 CCGTCGAGACA-----AAGGCACTTACGATCCGAGCGCGCTGACCGCGCGTC 3386543
;
Qy      212 AlaValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLys 231
;
Db      3386544 GACGTCATGTTCCGATTTGCTGACGCTGCGCAATCGCTGCTCTC----- 3386594
;
Qy      232 ValAlaSerTyrIleGluAspLysAspAsnIle-----LeuLeuTyrGly 246
;
Db      3386595 -----GAGTCGAGCGCATCGCGCGCGCGCCGCGCACGATCGCGCC 3386636
;
Qy      247 ThrAspIleGluPheIleGlyTyr----- 254
;
Db      3386637 TTCGACACCGAGTTGTTCCGCCATCGGTGTAAGAGGCCCAACTGCTGCAACGGGTA 3386696
;
Qy      255 ---ArgAspIleAlaGlyTyrArgMetSerValGluGlyLeuLeuValIleAspGlu 273
;
Db      3386697 TTGCGGAGCTTTACCGCGCGCGGTGTCGGGTGGGACCTGTGAGCATGATCGCGCGAC 3386756
;
Qy      274 LeuAsnSerGluLeuCybLeuProSerGluLeuLysHisSerGlyArgGluLeuTyrLeu 293
;
Db      3386757 -----GANTGTCGCGCGACCGCGCTGCAATTC 3386783
;
Qy      294 ArgThrSerSerTyrAlaProAspLysSerLeuArgIleTyrArgGluAspGluLysAsn 313
;
Db      3386784 CCACCCAGCTCTTGGGGTTCGCGCAAGGACTGCGCAGGTGTGG-----ACCGGTGCC 3386834
;
Qy      314 AlaArgLeuAsnMetLeuSerTyrAsnMetArgGlyGluLeuAla----- 328
;
Db      3386835 AAGGTGGCGGATGTGGTCAAGCTCAAC-----AGCGAAGTGGTCAATACCGGTTGACC 3386888
;
Qy      329 -----PheLeuAlaGluAsnSerAspAlaArgGlyTyrGluProLeuProGlu 344
;
Db      3386889 ACCATGCAACAAGCGGTGGCGACAGACGTCCTCGTACCGA-----CCGTCCTCGC 3386942
;
Qy      345 ArgArgLeu-----AspAlaPheArgAlaIleTyrAsnAspTyr 357
;
Db      3386943 GATCAAGTTGATGATGATCTCTGCGCAGAACCTGCTCACCGTTCACAGCGACTGG 3386999
;

```

```

;
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1

Alignment Scores:
Pred. No.: 3,32e+03 Length: 4411529
Score: 104.00 Matches: 76
Percent Similarity: 34.82% Conservative: 49
Best Local Similarity: 21.17% Mismatches: 128
Query Match: 5,53% Indels: 106
DB: 3 Gaps: 18

US-09-886-400a-4 (1-364) x US-09-103-840A-1 (1-4411529)

Qy      61 LeuValLysGlyLeuAlaSerAspLeuIleGlyThrSerTyrThrHis 80
;
Db      3390363 CTGCTGCGGCGCTGATGACCGCGACGAGTGAAGTCTGCTGCGGCCACTTGCCAC 3390422
;
Qy      81 AlaIleLeuProLeuLeuProLeuSerArgValGluAlaGlnValGlnArgAspArgGlu 100
;
Db      3390423 CCGTTCACCGCGCTGCTG-----GCACCGCGCGCTGCGCATTTCCGG 3390464
;
Qy      101 ValLysGluGluLeuPheGluValSer-----ProLysGlyPheTyr 114
;
Db      3390465 CTGCGGAAAGGCTCGCGCATCTCACTGCGCGTGGCGCACCGCCGAAAGGATCTGG 3390524
;
Qy      115 LeuProGluLeuAlaTyrAspProIleIleProAlaIleLeuLysAspAngLYTyrGlu 134
;
Db      3390525 GCACCCGAATGGCATAGCGCCCGGGGATGAGGTGCACTACGCCACCGCGGGGTCACT 3390584
;
Qy      135 TyrLeuPheAlaAspGlyGluAlaMetLeuPheSerLahIleuAsnSerAlaIle--- 153
;
Db      3390585 CACTTCATGCTGACGGCGCG-----TCGTCGACGAGGACACCGCGCTGGGC 3390632
;
Qy      154 LysProIle----- 156
;
Db      3390633 CGGCGGTGGGGAACCGATGCTGCGCTTCGCTGCGCACTTGCAGGTCACTACCGG 3390692
;
Qy      157 -----LysProLeuTyrPro---HisLeuIleLysAlaGlnArgGluLysArg 171
;
Db      3390693 GTGTGATCACGAAATCCGGCTACCCCGGGCAC-----GCCGCTTACCGCGAC----- 3390740
;
Qy      172 PheArgTyrIleSerTyrLeuLeuGlyLeuArgGluLeuArgLysAlaIleLysLeuVal 191
;
Db      3390741 TTCACACCTACGACACCTGACCGGACTCAACCGGCGAGGTGACCGGCGGTAACTGG 3390800
;
Qy      192 PheGluGlyLysValThrLeuLysAlaValLysAspIleGluAlaValProValTyrVal 211
;
Db      3390801 CCGTCGAGACA-----AAGCACCTTACGATCCGAGCGGCGTACCGCGCTTC 3390851
;
Qy      212 AlaValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLys 231
;
Db      3390852 GACGTCATGTTGCGGATTTGCTGACGCTGAGCGCAATCGGCTCTCTCC----- 3390902
;
Qy      223 ValAlaSerTyrIleGluAspLysAspAsnIle-----LeuLeuTyrGly 246
;
Db      3390903 -----GAGTCGAGCGCATCGCGCGCGCGCCGCGCCACGATGATCGCGCC 3390944
;
Qy      247 ThrAspIleGluPheIleGlyTyr----- 254
;
Db      3390945 TTGCAACACGAGTTGTTGGCCACTGTGTGACGAGGCGCCCAACTGCTGCAACGGGTA 3391004
;
Qy      255 ---ArgAspIleAlaGlyTyrArgMetSerValGluGlyLeuLeuValIleAspGlu 273
;
Db      3391005 TTGCGGAGCTTACCGCGCGCGGTGTCCGGGTGGGACCGCTGAGGAGATGCGTCCGCGAC 3391064
;

```

```

RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS

```

Qy 274 LeuAnSerGluLeuCyLeuProSerGluLeuLysSerGlyArgGluLeuTyrLeu 293
 Db 3391065 -----GGATTCGTGGGAGACCCGGTCAATTG 3391091
 Qy 294 ArgThrSerSerTrpAlaProAspLysSerLeuArgIleTrpArgGluAspGluLys 313
 Db 3391092 CCACCCAGACTCTTGGGTTCCGGCAGACGCTGCGAGGTTGG-----AGCGGTCC 3391142
 Qy 314 AlaArgLeuAsnMetLeuSerTyrAsnMetArgIleGluLeuAla----- 328
 Db 3391143 AAGTGCCGATCTGTGCTCAGCTCAAC-----ACCGAAGTGTGATACCGCGTTGACC 3391196
 Qy 329 -----PheLeuAlaGluAsnSerAspAlaArgIleTyrGluProLeuProGlu 344
 Db 3391197 ACCATCGACAAAGCGCTGGCCAGACGCTCCCTGACGGA-----CCGCTGCTCGC 3391250
 Qy 345 ArgArgLeu-----AspAlaPheArgAlaIleTyrAsnAspTrp 357
 Db 3391251 GATCAGCTGTGATCAGATCCTCGCGAGACCCCTGCTCACCCTGTCACGAGACTGG 3391307
 RESULT 10
 US-08-956-171E-324/c
 ; Sequence 324, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunach
 ; Gil H. Choi
 ; Patrick S. Dillon
 ; Craig A. Rosen
 ; Steven C. Barash
 ; Michael R. Fannon
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/956,171E
 ; FILING DATE: 20-Oct-1997
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/009,861
 ; FILING DATE: January 5, 1996
 ; APPLICATION NUMBER: 08/781,986
 ; FILING DATE: January 3, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mark J. Hyman
 ; REGISTRATION NUMBER: 46,789
 ; REFERENCE/DOCKET NUMBER: PB348P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (240) 314-1224
 ; TELEFAX: (301) 309-8439
 ; INFORMATION FOR SEQ ID NO: 324:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5030 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 324:
 US-08-956-171E-324
 Alignment Scores:
 Pred. No.: 0.0572 Length: 5030

Score: 103.00 Matches: 58
 Percent Similarity: 38.6%
 Best Local Similarity: 21.17% Mismatches: 48
 Query Match: 5.48% Indels: 90
 DB: 4 Gaps: 78
 US-09-886-400A-4 (1-364) x US-08-956-171E-324 (1-5030)
 Qy 107 GluValSerProLysGlyPheTrpLeuProGluLeuAlaTyrAspProIleIleProAla 126
 Db 2699 GAGATATCACT-----TATTATGATTCAATTA----- 2673
 Qy 127 IleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGlyLysAlaMetLeuPheSer 146
 Db 2672 CTCGTAAATATTATCTACACGCGATATCATTTAAACAGCAGAGAAAAATGTACGC 2613
 Qy 147 AlaHisLeuAsnSerAlaIleLysProIleLysProLeuTyrProHisLeuIleLysAla 166
 Db 2612 TCATTAGAGAAATAGCGATTCGTGTGTAAACTAATTCATTCTTATTAATAGTA 2553
 Qy 167 GluArgGluLysArgPhe-----ArgTyrIleSerTyrLeuLeu 179
 Db 2552 ATGAAAGATATAAAAGTTCAACAGTGTGATTAACAATAATTATTGAAGAAACACA 2493
 Qy 180 GlyLeuArgGluLeuArgLysAlaIle-----LysLeuValPheGluGly 194
 Db 2492 GAACCTTTTCGACATTCAGCGCTCTGTAGATAGAGCTAATAACATTAGATATAGGT 2433
 Qy 195 LysValThrLeuLysAlaValLysAspIleGlu----- 205
 Db 2432 AATGTAAACAATTATAGTGTTCGCCAATGTGAGAAAGTCGAAACGACTAGTAATTA 2373
 Qy 206 ---AlaValProValTrpValAlaValAsnThrAlaValMetLeuGlyIleGlyArg-- 223
 Db 2372 GCATCAATTCACATGATATCTTCAAGTAAATCCCTTCATTAGGTGACGAAACAATTG 2313
 Qy 224 LeuProLeuMetAsnProLysLysValAlaSerTrpIleGluAspLysAspAsnIleLeu 243
 Db 2312 CTGTATAGATAGTGTCCAAAGAGTGTGATGAGGTAAAGAGGATGATGCTTA 2253
 Qy 244 LeuTyrGlyThrAspIleGluPhe---IleGlyTyrArgAspIleAlaGlyTyrArgMet 262
 Db 2252 CTA-----ACAGTACACACTTTAGAGATGACACACCAATCATTAATGACACGAGTT 2199
 Qy 263 SerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCyLeuProSer 282
 Db 2198 AGACTAAGATATGATTAATATGTCATCC----- 2169
 Qy 283 GluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaProAspLys 302
 Db 2168 -----AAAACACCGACGATATTAAAGATGCT 2142
 Qy 303 ---SerLeuArgIleTrpArgGluAspGluLysAsnAlaArgLeuAsnMetLeuSerTyr 321
 Db 2141 TTCTCATTAAGAAATGTG-----GGCGTCTCATTTGAT---GTGCAATAT 2097
 Qy 322 AsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTyrGluPro 341
 Db 2096 AAT-----TTCTTGAAGAAAAACCA-----TGGAAACGA 2067
 Qy 342 LeuProGluLysArgLeuAspAlaPheArgAlaIleTyrAsn 355
 Db 2066 CTTGAAAGTCTAGT-----AAAGTAATTCCAAT 2037
 RESULT 11
 US-08-781-986A-324/c
 ; Sequence 324, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunach
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:

```

ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8504
INFORMATION FOR SEQ ID NO: 324:
SEQUENCE CHARACTERISTICS:
LENGTH: 5030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-324

Alignment Scores:
Pred. No.: 0.0572 Length: 5030
Score: 103.00 Matches: 58
Percent Similarity: 38.69% Conservative: 48
Best Local Similarity: 21.17% Mismatches: 90
Query Match: 5.48% Indels: 78
DB: 4 Gaps: 15

US-09-886-400a-4 (1-364) x US-08-781-986A-324 (1-5030)
QY 107 GluValSerProLysGlyPheTrpLeuProGluLeuAlaIleTyrAspProIleLeuProAla 126
Db 2699 GAGATATCACCT-----TATTATGATTCATTA----- 2673
QY 127 IleLeuLysAspAsnGlyTyrGluTyrLeuPheAlaAspGlyGluAlaMetLeuPheSer 146
Db 2672 CTGCTAAATTACTCTACACGCGATATCATTTTAAACAAGCAAGAAAAAATGTAACGC 2613
QY 147 AlaHisLeuAsnSerLaiIleLysProIleLysProLeuTyrProHisLeuIleLysAla 166
Db 2612 TCATTACGAGAAATGCGTATTCGTGGTGTAAACTAATATTCATTCAATTTAATGTA 2553
QY 167 GluArgGluLysArgPhe-----ArgTyrIleSerTyrLeuLeu 179
Db 2552 ATGAGAATAATAAAAAGTTTCAACAAGTGTGATTACACAACATAAATTTTGAAGAAACCA 2493
QY 180 GlyLeuArgGluLeuArgLysAlaIle-----LysLeuValPheGluGly 194
Db 2492 GAACCTTTGCAACATTGACCGCTCTAGAGTAGAGTACTTAAACATTAAGATATATAGGT 2433
QY 195 LysValThrLeuLysAlaValLysAspIleGlu----- 205
Db 2432 AATGTAAACAATTATATGCTTTCCAAATGTTGAGAAAAGTCCGAAACCAAGTATGAAATTA 2373
QY 206 ---AlaValProValIleTyrValAlaValAsnThrAlaValMetLeuGlyIleGlyArg--- 223
Db 2372 GCATCAATTCCAACGTATCTCAAGTAATAATGCTTCATTTAGTGTGAGAAACAATG 2313
QY 224 LeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspLysAspAsnIleLeu 243
Db 224 LeuProLeuMetAsnProLysLysValAlaSerTyrIleGluAspLysAspAsnIleLeu 243
```

```

Db 2312 CTTGATGAGTAGTACGCAAAAGGTGATGCTGAATGGGTTAAAAAGCAGATGATGCTTAA 2253
QY 244 LeuTyrGlyThrAspIleGluPhe---IleGlyTyrArgAspIleAlaGlyTyrArgMet 262
Db 2252 CTA-----ACAGATACCAACCTTTAGAGATGACACACCAATCATTTATTACTACACAGATT 2199
QY 263 SerValGluGlyLeuLeuGluValIleAspGluLeuAsnSerGluLeuCysLeuProSer 282
Db 2198 AGAAGTAAAGATATGATTAATATCGCATCC----- 2169
QY 283 GluLeuLysHisSerGlyArgGluLeuTyrLeuArgThrSerSerTrpAlaProAspLys 302
Db 2168 -----AAAACGCGACGCTATTTAAGATGAT 2142
QY 303 ---SerLeuArgIleTrpArgGluAspGluGlyAsnAlaArgLeuAsnMetLeuSerTyr 321
Db 2141 TTCTCACTGAAGATGTGG-----GGCGGTGTACATTTGAT---GTGCGATAT 2097
QY 322 AsnMetArgGlyGluLeuAlaPheLeuAlaGluAsnSerAspAlaArgGlyTrpGluPro 341
Db 2096 AAT-----TTCTTGAAAGAAAACCA-----TGGGAACGA 2067
QY 342 LeuProGluArgArgLeuAspAlaPheArgAlaIleTyrAsn 355
Db 2066 CTGGAACGTCTACGT-----AAAGCTATTCCAAT 2037

RESULT 12
US-08-270-013B-1
Sequence 1, Application US/08270013B
Patent No. 5686294
GENERAL INFORMATION:
APPLICANT: Sogabe et al.
TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
NUMBER OF INVENTION: 2
NUMBER OF SEQUENCES: DEHYDROGENASE ACTIVITY
CORRESPONDENCE ADDRESS:
ADDRESSER: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 61601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/270,013B
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 164701/1993
FILING DATE: 02-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Robert F.
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 62321
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: (25)3533
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus stearothermophilus
```

```

; STRAIN: ATCC12016
US-08-270-013B-1

Alignment Scores:
Pred. No.: 0.0133 Length: 1912
Score: 102.50 Matches: 64
Percent Similarity: 37.54% Conservative: 46
Best Local Similarity: 21.84% Mismatches: 92
Query Match: 5.46% Indels: 92
DB: 1 Gaps: 14

US-09-886-400A-4 (1-364) x US-08-270-013B-1 (1-1912)

QY 10 LeuGlnTyraLagluIleProlysSerGluIleProlys----- 22
DB 1070 ATGACGTAACGCTTTAAGAAATCCGATTCGGAACAAACCGGCTATCGCCACATCG 1129
QY 23 ---ValIleGluYsaIaTyriLeProValIleGluThrLeuIleYsGluGluIlePro 41
DB 1130 GCGGCTTGATACGCGCGC-----TTCCGACGTTGCTGCGCAGAG----- 1174
QY 42 PheGlyLeuAenIleThrcIyTyriThreLeuYsPheLeuProYsaPilleIleAspleu 61
DB 1175 -----CTGAACATTTCG-----GTAAAGATGCTACTGCGGCTTT 1207
QY 62 ValYsGlyGlyIleAlaSerAspleuIleGluIleIleGlyThrcTyThriIleAla 81
DB 1208 GTTTTAGCGCGCATGCGGATGACATGCTGCGCTGCTCGCTACTGTCGCGCGCGC 1267
QY 82 Ile-----LeuProLeuLeuProLeuSerArgValIleGluIleAlaGlnValIleAspArg 99
DB 1268 ATTCGCGTCGAAAAATCATTCGGAAGATCGTTTGAAGCGCATGTTGAGCGAGCGCGC 1327
QY 100 GluValYsGluGluLeuPheGluValSerProYsGly-----PheTyLeuProGlu 117
DB 1328 AAAGCGCGCGGTAATGCTCAACCTGCTCGCAACGCGCAGCGCTACTACGACCGCGC 1387
QY 118 LeuIaTyraPProIleIleProAlaIleLeuYsAspAn----- 131
DB 1388 GCCTGCTTTCGAAATGCTGAAACGATTTTGAAGAACAGCGCGCATTTTGGCGCGC 1447
QY 132 -----GlyTyriGluTyriLeuPheAlaAspGlyGluIleAlaMet 143
DB 1448 ATGCGCTAATCTGAAGCGGATACGCGCTATGAGGCAATTATTGGCGTCCGACATC 1507
QY 144 LeuPheSerAlaHISLeuAsnSerAlaIleYs----- 154
DB 1508 CTCGCGCGCAACGCGCATCGAAGAAAGTATCGAGCTCGAGCTGACCGAAGAGAAAAACG 1567
QY 155 -----ProIleYs----- 157
DB 1568 CGCTGCGCAAAATCCGTCGAATCCGTTAAAAATGTCATCGCATCGTGAATAGCGCGAG 1627
QY 158 -----ProLeuTyriProHISLeuIleYsAlaGlnArgGluYsArgPhe----- 172
DB 1628 GCAAAATATCCGCGATTCGCGGATTTTGTTCACACACTCAATGAAGCGCTTTCTAGAC 1687
QY 173 ---ArgTyriLeSerTyriLeuLeuGlyLeuArgLeuArgYsaIaIle----- 188
DB 1688 AACGAAGGGGTGGGAC-ATGTTGAAAAACGAAGCTCGGAGAGCGCATCGGGAAATT 1746
QY 189 -----LyLeuValPheGluGlyYsValThrLeuYsAlaValIleYsaPille 204
DB 1747 CAAGCGGGGAAAAAGCTGTTCCAAAGCCGCATC-----GAAGACAAAGACTTG 1797
QY 205 -----GluIleValProValITrProValAlaValIleAsnThraIa 216
DB 1798 CTTCTTATCTTGAGCTGACGAGATGATCCATCGCTCTATATCCAGCATGATTATGCT 1857
QY 217 ValMetLeuGlyIleGlyArgLeuProLeuMetAsnPro 229
DB 1858 TCACAGACGCGCTTTGAAAA---CGGCTGTCGCGCGC 1893

```

```

RESULT 13
US-08-838-418-1
; Sequence 1, Application US/08838418
; Patent No.: 574342
; GENERAL INFORMATION:
; APPLICANT: Sogabe et al.
; TITLE OF INVENTION: PROTEIN HAVING HEAT-RESISTANT MALATE
; TITLE OF INVENTION: DEHYDROGENASE ACTIVITY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,418
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/270,013
; FILING DATE: 01-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 164701/1993
; FILING DATE: 02-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoover, Allen E.
; REGISTRATION NUMBER: 37354
; REFERENCE/DOCKET NUMBER: 78339
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: (25) 3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1912 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus stearothermophilus
; STRAIN: ATCC12016
US-08-838-418-1

Alignment Scores:
Pred. No.: 0.0133 Length: 1912
Score: 102.50 Matches: 64
Percent Similarity: 37.54% Conservative: 46
Best Local Similarity: 21.84% Mismatches: 92
Query Match: 5.46% Indels: 92
DB: 1 Gaps: 14

US-09-886-400A-4 (1-364) x US-08-838-418-1 (1-1912)

QY 10 LeuGlnTyraLagluIleProlysSerGluIleProlys----- 22
DB 1070 ATGACGTAACGCTTTAAGAAATCCGATTCGGAACAAACCGGCTATCGCCACATCG 1129
QY 23 ---ValIleGluYsaIaTyriLeProValIleGluThrLeuIleYsGluGluIlePro 41
DB 1130 GCGGCTTGATACGCGCGC-----TTCCGACGTTGCTGCGCAGAG----- 1174
QY 42 PheGlyLeuAenIleThrcIyTyriThreLeuYsPheLeuProYsaPilleIleAspleu 61

```

```
Db      1175 -----CTGAACATTTCG-----CTAAAGATGTCACGTGGGTTT 1207
Qy      62  Vallylgllyllylealaserpleuilegluilelythrserlythrhisala 81
Db      1208  GTTTTAGGCGCCGATGCGGATGACATGTCGCCCTCGTCGCTACCTGTCGCGCGGCG 1267
Qy      82  Ile-----LeuProleuLeuProleuSerArgValGluValGlnValGlnArgAspArg 99
Db      1268  ATTCCGCTCGAAAAATCTCATCTCCGAAAAGATCGTTTGGACCCCATCGTTGACCGGACGCGC 1327
Qy      100  GluVallylgllylueuphegluValSerProlylsgly-----PheTrpLeuProGlu 117
Db      1328  AAGCGCGCGCGTGAATCGTCAACCTGCTCGGACGCGGACGCGCTACTACGACGCGCGCC 1387
Qy      118  LeuAlaTyraSprIleIleIleProAlaIleLeuLysAspAsn----- 131
Db      1388  GCCTCGCTGTGCAAAATGTCGAAGCGCATTTTGAAGACCGCGCGCATTTTTCGCGCGC 1447
Qy      132  -----GlyTyrglyTyrrheupheAlaAspGlyGluAlaMet 143
Db      1448  ATCGCTACTCTGAAGCGCAATACGCTATGAAGGCAATTATTGGCGCTGCCGACGATC 1507
Qy      144  LeuPheSerAlaHisleuAsnSerAlaIleLys----- 154
Db      1508  CTCGCGCGGCAACGCGCATCGAAGAAAGTGCAGCTGAGCTGACCGAAGGAAAGG 1567
Qy      155  -----ProIleLys----- 157
Db      1568  CGCTCGCAAAATCCGTCGAATCCGTTAAATAATGTCATGCGCATCGTGAATAGCGCGAG 1627
Qy      158  -----ProleuTyrrProHisleuIleLysAlaGlnArgGlyLysArgPhe----- 172
Db      1628  GCAAAATTCGCGCATTTGCCGATTTTTCCTCCACAGTCAATGAAGACCTTTCTTAGAC 1687
Qy      173  ---ArgTyrrIleSerTyrrleuLeuArgGlyLeuArgGlyLysAlaIle----- 188
Db      1688  AACGAAGGGGTGGGAAC-ATGTTGAAAAAAGCAAGCTCGGAGACCGATCGGGGAAAT 1746
Qy      189  -----LysLeuValPheGluGlyLysValThrLeuLysAlaValLysAspIle 204
Db      1747  CAAGCGGGGAAAAAGTCGTGTTCCAAAGCCGCATC-----GAAGAAGAAAGCTTG 1797
Qy      205  -----GluAlaValProValTrpValAlaValAsnThrAla 216
Db      1798  CTTCTTATCTTGCGGTGACGATGATGCCATCCGCTATATCCAGCATGATATGCT 1857
Qy      217  ValMetLeuGlyIleGlyArgLeuProLeuMetAsnPro 229
Db      1858  TCACAGACGCGGTTTGGAAAA---CCGCTCGTGCCTCGCG 1893

RESULT 14
US-09-710-279-2149
; Sequence 2149, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2149
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-2149
```

```
Alignment Scores:
Pred. No.: 0.00693 Length: 1083
Score: 101.50 Matches: 57
Percent Similarity: 36.45% Conservative: 52
Best Local Similarity: 19.06% Mismatch: 117
Query Match: 5.40% Indels: 73
DB: 4 Gaps: 12

US-09-886-400a-4 (1-364) x US-09-710-279-2149 (1-1083)
Qy      4  LeuValPheHisGlyAsnLeuGlnTyrrAlaGluIleProLysSerGluIleProLysVal 23
Db      181  CTTATTTTCAAGGA-----GTAAGACGTATC 207
Qy      24  IleGluLysAlaTyrrIleProValIleGluThrLeuIleLysGluGluIleProPheGly 43
Db      208  GTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 261
Qy      44  LeuAsnIleThrGlyTyrrThrLeuLysPheLeuProLysAspIleIleAspLeuValLys 63
Db      262  CTATATGCACTTACACATATCTTGATGTAATCTTAAGGTGTCATGCAATGATGATGATGATG 321
Qy      64  GlyGlyIleAlaserAspLeuIleGluIleIleGlyThr-----SerTyrrThrHisala 81
Db      322  GATCAAAATGCTTTAGAGAAACATATCATGATGATTAATACAAATAGTCATATTTATTAACA 381
Qy      82  IleLeuProleuLeuProleuSerArgValGluAlaGlnValGlnArgAspArgGluVal 101
Db      382  GTTCAAACTTTTATACCTTAATAATTAATTAATGA-----GATTTTC 420
Qy      102  LysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTyrrAsp 121
Db      421  AAGACAGTTTAAACGAACCT----- 441
Qy      122  ProIleIleProAlaIleLeuLysAspAsnGlyTyrrGlyTyrrheupheAlaAspGlyGlu 141
Db      442  -----CGATTACTGAAGAAAGTAATTACGAATATGTTCTTTGAAAGTGAA 489
Qy      142  AlaMetLeuPheSerAlaHisleuAsnSerAlaIleLysProIleLysProleuTyrrPro 161
Db      490  GGT-----AAAGGCAATTTTAAACGATGATGATGATGATGATGATGATGATGATGATGAT 528
Qy      162  HisLeuIleLysValIleGlnArgLysArgPheArgTyrrIleSerTyrrheupheGlyLeu 181
Db      529  TATATGGAAGATTAGATAGT-----ATCGAATATGTT----- 561
Qy      182  ArgGluLeuArgLysAlaIleLysLeuValPhe-----GluGlyLysVal 196
Db      562  -----GATGAATTAATAAATTGAGTTTATGATGAAGACATGAAATTAAGAAATTA 609
Qy      197  ThrLeuLysAlaValLysAspIleGlu-----AlaValProVal-----TrpValAla 212
Db      610  ACTTAAACGTGCTATTTTGAATATATCAACCATACGAAACACGATTTTGAATTTTAA 669
Qy      213  ValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLeuMetAsnProLysVal 232
Db      670  ATGAACAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
Qy      233  AlaSerTrpIleGluAspLysAspAsnIleLeuLeuIleLeuLysGlyTyrrAspIleGluPhe 252
Db      730  TTAGATGAATTTTCTGAATATGCAAAAAACAGCTCAATATACCGAGCGTACGATATTA 789
Qy      253  GlyTyrrArgAspIleAlaGlyTyrrArgMetSerValGluGlyLeuLeuGluVal----- 270
Db      790  GGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
Qy      271  -----IleAspGluLeuAsnSerGluLeuCysLeuProSerGluLeuLysHis 286
Db      850  GAGTATTAAGCTTAGCAACTTGAGCAGATGTTTGTGATCGTGATATTAACAC 906

RESULT 15
US-09-710-279-2105
```

```

; Sequence 2105, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU348005
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2105
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-2105

```

```

Alignment Scores:
Pred. No.: 0.00745 Length: 1131
Score: 101.50 Matches: 57
Percent Similarity: 36.45% Conservative: 52
Best Local Similarity: 19.06% Mismatches: 117
Query Match: 5.40% Indels: 73
DB: 4 Gaps: 12

```

US-09-886-400a-4 (1-364) x US-09-710-279-2105 (1-1131)

```

QY 4 LeuValPheHisGlyAsnLeuGlnTrpAlaGluIleProLysSerGluIleProLysVal 23
DB 139 CTATTTTCAAGGA-----GTAACGATAC 165
QY 24 IleGluValAlaTrpIleProValIleGluTrpLeuIleLysGluGluIleProPheGly 43
DB 166 GTTGAGAGTGGATGATGATATATTCGTAACCTTATCCAAATATATC-----AAT 219
QY 44 LeuAsnIleTrpGlyTrpLeuLysPheLeuProLysAspIleIleAspLeuValLys 63
DB 220 CTATAGCATTAACACATATCTGATGTAAATCCTAAAGGTCATCGAATGTTACG 279
QY 64 GlyGlyIleAlaSerAspLeuIleGluIleIleGlyThr-----SerTrpThrHisAla 81
DB 280 GATCAAAATGAGTTAGAGAACATATCAATGATTAATACAAATAGCTCATATTATTACAA 339
QY 82 IleLeuProLeuLeuSerArgValGluAlaGluValGlnArgAspArgGluVal 101
DB 340 GTTCAAACTTTTATATCTTAAATATATATGAA-----GATTTC 378
QY 102 LysGluGluLeuPheGluValSerProLysGlyPheTrpLeuProGluLeuAlaTrpAsp 121
DB 379 AAAGACAGTTTAAACGAACTT----- 399
QY 122 ProIleIleProAlaIleLeuLysAspAsnGlyTrpGluTrpLeuPheAlaAspGlyGlu 141
DB 400 -----GGATTAGCTAAAGAGTAATTCGAATATGTTCTTGAAAGTGA 447
QY 142 AlaMetLeuPheSerAlaHisLeuAsnSerAlaIleLysProIleLysProLeuTrpPro 161
DB 448 GGT-----AAAGGCAATTTAAACCGATGAGTGAATGCTCT 486
QY 162 HisLeuIleLysAlaGlnArgGluLysArgPheArgTrpIleSerTrpLeuLeuGlyLeu 181
DB 487 TATATAGGGAAGTATGATAGT-----ATCGAATATGTT----- 519
QY 182 ArgGluLeuArgLysAlaIleLysLeuValPhe-----GluGlyLysVal 196
DB 520 -----GATGAATATAAACTTGATTTATGATAAAGACATGAATTAAGAAATA 567
QY 197 ThrLeuLysAlaValLysAspIleGlu-----AlaValProVal-----TrpValAla 212
DB 568 ACTAAACGTGCTATTTTATGATTAATCACCCTATACGAAACACCGATTTTATTTATTA 627
QY 213 ValAsnThrAlaValMetLeuGlyIleGlyArgLeuProLysMetAspProLysVal 232
DB 628 ATGAACAAAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
QY 233 AlaSerTrpIleGluAspLysAspAsnIleLeuLeuTrpGlyTrpAspIleGluPheIle 252
DB 688 TTAAGTGAATTTTCTGAATATGCCCCAAAAACAGCTCAATATACGAGGTACGATATACA 747
QY 253 GlyTrpArgAspIleAlaGlyTrpArgMetSerValGluGlyLeuLeuGluVal----- 270
DB 748 GGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
QY 271 -----IleAspGluLeuAsnSerGluLeuGlyLeuProSerGluLeuLysHis 286
DB 808 GAGTATTAAGCTACGCAACTTGAGAGCATGTTTGTACTGATGATTAACAC 864

```

Search completed: January 28, 2005, 06:04:35
Job time : 3883 secs

This Page Blank (uspto)